

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа Инженерная школа информационных технологий и робототехники
Направление подготовки 09.04.01 Информатика и вычислительная техника
Отделение школы (НОЦ) Отделение информационных технологий

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Разработка алгоритма обнаружения лиц в защитной маске на основе методов компьютерного зрения и искусственного интеллекта

УДК 004.421:004.932.72:004.8

Студент

Группа	ФИО	Подпись	Дата
8BM93	Султонов Ансорхон Азизхонович		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Друки А.А.	К.Т.Н.		

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Верховская М.В.	К.Э.Н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ООД ШБИП	Федоренко О.Ю.	Д.М.Н.		

По разделу на иностранном языке

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель ОИЯ ШБИП	Розанова Я.В.	—		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОИТ ИШИТР	Спицын В.Г.	Д.Т.Н.		

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

по направлению 09.04.01 Информатика и вычислительная техника

Код компетенции	Наименование компетенции
	Универсальные компетенции
УК(У)-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий
УК(У)-2	Способен управлять проектом на всех этапах его жизненного цикла
УК(У)-3	Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели
УК(У)-4	Способен применять современные коммуникативные технологии, в том числе на иностранном (-ых) языке (-ах), для академического и профессионального взаимодействия
УК(У)-5	Способен анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия
УК(У)-6	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки
	Общепрофессиональные компетенции
ОПК(У)-1	Способен самостоятельно приобретать, развивать и применять математические, естественно-научные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
ОПК(У)-2	Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач
ОПК(У)-3	Способен анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями
ОПК(У)-4	Способен применять на практике новые научные принципы и методы исследований
ОПК(У)-5	Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем
ОПК(У)-6	Способен разрабатывать компоненты программно-аппаратных комплексов обработки информации и автоматизированного проектирования
ОПК(У)-7	Способен адаптировать зарубежные комплексы обработки информации и автоматизированного проектирования к нуждам отечественных предприятий
ОПК(У)-8	Способен осуществлять эффективное управление разработкой программных средств и проектов
	Профессиональные компетенции
ПК(У)-1	Способен разрабатывать и администрировать системы управления базами данных

ПК(У)-2	Способен проектировать сложные пользовательские интерфейсы
ПК(У)-3	Способен управлять процессами и проектами по созданию (модификации) информационных ресурсов
ПК(У)-4	Способен осуществлять руководство разработкой комплексных проектов на всех стадиях и этапах выполнения работ
ПК(У)-5	Способен проектировать и организовывать учебный процесс по образовательным программам с использованием современных образовательных технологий

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа Инженерная школа информационных технологий и робототехники
Направление подготовки 09.04.01 Информатика и вычислительная техника
Отделение школы (НОЦ) Отделение информационных технологий

УТВЕРЖДАЮ:
Руководитель ООП

(Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации
(бакалаврской работы, дипломного проекта/работы, магистерской диссертации)

Студенту:

Группа	ФИО
8BM93	Султонову Ансорхону Азизхоновичу

Тема работы:

Разработка алгоритма обнаружения лиц в защитной маске на основе методов компьютерного зрения и искусственного интеллекта
Утверждена приказом директора (дата, номер)

Срок сдачи студентом выполненной работы:	
--	--

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

<p>Исходные данные к работе <i>(наименование объекта исследования или проектирования; производительность или нагрузка; режим работы (непрерывный, периодический, циклический и т. д.); вид сырья или материал изделия; требования к продукту, изделию или процессу; особые требования к особенностям функционирования (эксплуатации) объекта или изделия в плане безопасности эксплуатации, влияния на окружающую среду, энергозатратам; экономический анализ и т. д.).</i></p>	<p>Объектом исследования являются нейронные сети. Необходимо ее обучить и оценить способность распознавать лица, обнаружить маски на лице и обозначать их в кадре.</p>
<p>Перечень подлежащих исследованию, проектированию и разработке вопросов <i>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i></p>	<ul style="list-style-type: none"> — Обзор литературных источников. — Поиск или формирование выборок данных. — Сконфигурировать параметры нейронной сети. — Обучение нейронных сетей. — Сравнительный анализ результатов работы обученных нейронных сетей. — Социальная ответственность. — Финансовый менеджмент, ресурсоэффективность и ресурсосбережение. — Заключение.

Перечень графического материала (с точным указанием обязательных чертежей)	
--	--

Консультанты по разделам выпускной квалификационной работы (с указанием разделов)

Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Верховская М.В., Доцент ОСГН ШБИП, к.э.н.
Социальная ответственность	Федоренко О.Ю., Профессор ООД ШБИП, д.м.н.
Английский язык	Розанова Я.В., Старший преподаватель ОИЯ ШБИП

Названия разделов, которые должны быть написаны на русском и иностранном языках:
Анализ предметной области; Машинное обучение; Глубокое обучение; Переобучение нейронных сетей и методы предотвращения переобучения; Свёрточные нейронные сети; Методы обнаружения лиц;

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	
---	--

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Друки А.А.	К.Т.Н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8BM93	Султонов А.А.		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И РЕСУРСОСБЕРЕЖЕНИЕ»

Студенту:

Группа	ФИО
8BM93	Султонову Ансорхону Азизхоновичу

Школа	ИШИТР	Отделение школы (НОЦ)	ОИТ
Уровень образования	магистратура	Направление/специальность	09.04.01 «Информатика и вычислительная техника»

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Оклад инженера – 22 695,68 руб. Оклад руководителя – 35 111,5 руб.
2. Нормы и нормативы расходования ресурсов	Районный коэффициент 30%; Коэффициент дополнительной заработной платы 12%; Накладные расходы 16%.
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Коэффициент отчислений на уплату во внебюджетные фонды 30%

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого и инновационного потенциала НТИ	Анализ перспективности технических решений посредством Quad-анализа SWOT-анализ. Оценка готовности научно-исследовательского проекта к коммерциализации
2. Разработка устава научно-технического проекта	Определение цели научно-исследовательского проекта, требований к проекту, описание заинтересованных стороны проекта, рабочей группы
3. Планирование процесса управления НТИ: структура и график поведения, бюджет, риски и организация закупок	Планирование этапов работы, определение календарного графика проведения исследования. Определение рисков научно-исследовательского проекта, оценка вероятности риска и потерь. Расчет бюджета затрат на проведение исследования
4. Определение ресурсной, финансовой, экономической эффективности	Описание потенциального эффекта научно-исследовательского проекта.

Перечень графического материала (с точным указанием обязательных чертежей):

1. Оценочная карта технологии Quad
2. Матрица SWOT
3. Диаграмма Ганта
4. Бюджет затрат
5. Потенциальные риски

Дата выдачи задания для раздела по линейному графику	
--	--

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Верховская М.В.	Кандидат экономических наук		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8BM93	Султонов А.А.		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
8BM93	Султонову Ансорхону Азизхоновичу

Школа	Инженерная школа информационных технологий и робототехники	Отделение (НОЦ)	Информационных технологий
Уровень образования	Магистратура	Направление/специальность	09.04.01 Информатика и вычислительная техника

Тема ВКР:

Разработка алгоритма обнаружения лиц в защитной маске на основе методов компьютерного зрения и искусственного интеллекта	
Исходные данные к разделу «Социальная ответственность»:	
1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	Алгоритмы обнаружения лиц в защитной маске используются для решения разных задач, таких как: распознавание, обнаружение и детектирование лиц. Разработанный алгоритм может быть востребован в системах автоматической идентификации и распознавании людей. Рабочее место: площадь помещения составляет 45м ² , есть искусственное и естественное освещение, а также отопление и вентиляция, присутствует рабочий стол и персональный компьютер.
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
1. Правовые и организационные вопросы обеспечения безопасности: <ul style="list-style-type: none"> – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны. 	Мероприятия по защите от вредных факторов согласно нормативным документам: <ul style="list-style-type: none"> – Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ (ред. от 09.03.2021). – СанПиН 1.2.3685-21 Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания. – ГОСТ 12.0.003-2015 Опасные и вредные производственные факторы. – СП 52.13330.2016 Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95. – СН 2.2.4/2.1.8.562-96. Шум на рабочих местах, в помещениях жилых,

	<p>общественных зданий и на территории жилой застройки.</p> <ul style="list-style-type: none"> – ГОСТ 12.1.003-2014 ССБТ. Шум. Общие требования безопасности. – ТОИ Р-45-084-01 Типовая инструкция по охране труда при работе на персональном компьютере. – ГОСТ 12.1.030-81 Система стандартов безопасности труда (ССБТ). Электробезопасность. Защитное заземление. Зануление. – ГОСТ 12.1.038-82 Система стандартов безопасности труда (ССБТ). Электробезопасность. Предельно допустимые значения напряжений прикосновения и токов. – ГОСТ 26522-85 Короткие замыкания в электроустановках – ГОСТ Р 12.1.019-2017 ССБТ Электробезопасность. Общие требования и номенклатура видов защиты. – ГОСТ 12.1.004-91 Система стандартов безопасности труда (ССБТ). Пожарная безопасность. – ГОСТ 12.2.032-78 Система стандартов безопасности труда (ССБТ). Рабочее место при выполнении работ сидя.
<p>2. Производственная безопасность: 2.1. Анализ выявленных вредных и опасных факторов 2.2. Обоснование мероприятий по снижению воздействия</p>	<p>Рабочий процесс проводится в специализированном помещении, где могут быть такие вредные факторы как:</p> <ul style="list-style-type: none"> – недостаточная освещенность рабочей зоны; – повышенный уровень шума и вибрации; – отклонение параметров микроклимата в помещении; – повышенный уровень излучения электромагнитных полей. <p>Психофизические факторы:</p> <ul style="list-style-type: none"> – повышенная нагрузка на органы зрения; – длительные статические нагрузки; – нервно-эмоциональное напряжение. <p>Опасные факторы:</p> <ul style="list-style-type: none"> – поражение электрическим током; – статическое электричество; – короткое замыкание.
<p>3. Экологическая безопасность:</p>	<p>Влияние объекта исследования в процессе разработки и применения на литосферу происходит при утилизации комплектующих ПК и периферийных устройств (принтеры, МФУ, веб-камеры, наушники, колонки,</p>

	<p>телефоны), аккумуляторных батареек, люминесцентных ламп, макулатуры.</p> <p>Мероприятия по защите окружающей среды согласно нормативным документам:</p> <ul style="list-style-type: none"> – ГОСТ 17.4.3.04-85 Охрана природы (ССОП). Почвы. Общие требования к контролю и охране от загрязнения – ГОСТ Р 53692-2009 Ресурсосбережение. Обращение с отходами.
4. Безопасность в чрезвычайных ситуациях:	<p>Вероятные ЧС, инициируемые объектом исследования и возникающие на рабочем месте:</p> <ul style="list-style-type: none"> – пожар; – социальные чрезвычайные ситуации. <p>Наиболее типичные ЧС: пожар.</p>

Дата выдачи задания для раздела по линейному графику	01.03.2021
---	-------------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ООД ШБИП	Федоренко О.Ю.	Доктор медицинских наук		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8BM93	Султонов А.А.		

Реферат

Выпускная квалификационная работа содержит пояснительную записку на 122 листах, включает 39 рисунков, 20 таблиц, 38 источник литературы, 2 приложение.

Ключевые слова: компьютерное зрение, распознавание образов, свёрточные нейронные сети, машинное обучение.

Объектом исследования являются методы компьютерного зрения для обнаружения лиц, и, нейронные сети для распознавания маски на лице.

Цель работы: разработка алгоритмов и программных средств обнаружения лиц на видео-последовательности и выявления признаков наличия защитной медицинской маски на лице.

Задачами исследование является:

- 1) Провести аналитический обзор существующих алгоритмов и систем распознавания образов.
- 2) Поиск и анализ датасетов.
- 3) Разработать алгоритм обнаружения лиц и выявления признаков наличия защитной медицинской маски на лице.
- 4) Провести тестирование разработанной программной системы.

В результате проведённых экспериментов в соответствии с поставленными задачами были проведены исследования различных алгоритмов обнаружения лиц и распознавание объектов. Был проведен анализ актуальности разрабатываемого проекта, анализ наиболее распространённого набора данных изображении лиц в защитных масках, исследованы и изучены различные эффективные методы и подходы, позволяющие определить правильность получаемых результатов, а также была спроектирована система распознавания лиц и обнаружения защитной маски на лице в видеопотоке.

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ, СОКРАЩЕНИЯ

CNN – convolutional neural network.

ЗП – заработная плата.

ПР – прибыль.

РФ — Российская Федерация.

ГОСТ — Государственный стандарт.

СНиП – Строительные нормы и правила.

СН – строительные нормы.

СП – строительные правила.

СанПиН – Санитарные правила, нормы, правила и нормы, гигиенические норматив

Содержание

Введение	14
1. Анализ предметной области	15
1.1. Описание предметной области.....	15
1.2. Машинное обучение.....	16
1.3. Глубокое обучение.....	16
1.4. Переобучение нейронных сетей и методы предотвращения переобучения.....	18
1.5. Свёрточные нейронные сети.....	19
1.6. Системы распознавания лиц.....	23
1.7. Методы обнаружения лиц.....	25
1.8. Метод Виолы-Джоса.....	28
1.9. Методы предварительной обработки изображений	35
1.10. Методы распознавания лиц.....	36
2. Разработка системы распознавания лиц в защитной маске.....	37
2.1. OpenCV	37
2.2. Dlib.....	39
2.3. TensorFlow.....	40
2.4. NumPy	40
2.5. Используемые архитектуры нейронных сетей.....	41
2.6. Используемые наборы данных	43
2.7. Обработки кадров видеопотока.....	45
3. Результаты тестирования модели	49
3.1. Поворот вокруг вертикальной оси.....	53
3.2. Поворот вокруг оси Z	54
3.3. Критерии принятия решения модели.....	54
4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	56
4.1. Предпроектный анализ	56
4.2. Технология QuaD.....	57
4.3. SWOT-анализ	59
4.4. Оценка готовности проекта к коммерциализации	64
4.5. Планирование управления научно-техническим проектом	67
4.5.1. План проекта.....	67
4.5.2. Бюджет научного исследования	69

4.5.2.1.	Специальное оборудование для научных работ	69
4.5.2.2.	Основная заработная плата.....	70
4.5.2.3.	Дополнительная заработная плата научно-производственного персонала	72
4.5.2.4.	Отчисления на социальные нужды.....	73
4.5.2.5.	Накладные расходы	73
4.5.2.6.	Формирование бюджета затрат научно-исследовательского проекта.....	74
4.6.	Риски научно-исследовательского проекта	74
4.7.	Выводы по разделу «финансовый менеджмент, ресурсоэффективность и ресурсосбережение».....	75
5.	Социальная ответственность	77
	Введение	77
5.1.	Правовые и организационные вопросы обеспечения безопасности	78
5.2.	Производственная безопасность.....	80
5.2.1.	Отклонение показателей микроклимата от нормы.....	81
5.2.2.	Недостаточная освещенность рабочей зоны	83
5.2.3.	Производственные шумы	87
5.2.4.	Психофизиологические факторы	88
5.2.5.	Опасность поражения электрическим током	88
5.2.6.	Опасность возникновения пожара.....	89
5.3.	Мероприятия и по снижению воздействия вредных производственных факторов.....	90
5.4.	Экологическая безопасность	91
5.5.	Безопасность в чрезвычайных ситуациях	92
5.6.	Выводы по разделу «Социальная ответственность».....	94
	Заключение	95
	Список использованных источников	96
	Приложение А	100
	Приложение Б. Исходный код процесса обучения и проверки моделей	118

Ведение

Новое заболевание, под названием коронавирус (COVID-19), серьезно повлияло на мир. К концу ноября 2020 года, согласно информации всемирной организации здравоохранения (ВОЗ), общее количество новых случаев коронавируса уже превысило 60 миллионов. На сегодняшний день, чтобы ограничить распространение болезни, использование масок становятся обычным явлением в общественных местах. Кроме того, многие поставщики государственных услуг требуют, чтобы клиенты носили маски в соответствии с заранее определенными правилами (например, закрывали рот и нос) при посещении определенных мест или территории. Данная ситуация во всем мире подтолкнула на исследования автоматизированных (основанных на компьютерном зрении) методов обнаружения масок, которые могут помочь контролировать общественное поведение и предотвратить распространении пандемии COVID-19.

Для распознавания определенных групп лиц и предоставления им доступа, на предприятиях устанавливают электронные турникеты с идентификацией отпечатков пальцев или со считывателями карт, но на сегодняшний день, в связи с бурным развитием биометрических технологий, компании все чаще переходят на другие методы распознавания, более точные и удобные. Например, распознавание лиц в видеопотоке в режиме реального времени.

Наряду с развитием методов распознавания и стремительного роста применения биометрических систем, методы идентификации по лицу проявляют свои недостатки и становятся уязвимыми.

В эпидемических ситуациях, таких как пандемия коронавируса (COVID-19), защитные медицинские маски стали необходимостью для людей, находящихся в общественных местах. Маски являются защитным средством,

которое используют люди в повседневной жизни для защиты от коронавируса. Такая ситуация делает обычную технологию контроля доступа неэффективной во многих случаях, а именно таких как контроль доступа в помещения, контроль доступа с использованием идентификации по лицу, подсчет посещаемости, проверка безопасности на вокзалах и т.д. В связи с большим спросом к технологиям распознавания и идентификации, особенно в сегодняшние дни, возникла большая необходимость улучшения методов распознавания для существующих способов технологии распознавания лиц, применительно к лицам, использующим защитные маски.

В связи с этим была определена цель данной работы – разработка алгоритмов и программных средств обнаружения лиц на видео-последовательности и выявления признаков наличия защитной медицинской маски на лице. Для достижения поставленной цели необходимо решить следующие задачи:

- 5) Провести аналитический обзор существующих алгоритмов и систем распознавания образов.
- 6) Поиск и анализ датасетов.
- 7) Разработать алгоритм обнаружения лиц и выявления признаков наличия защитной медицинской маски на лице.
- 8) Провести тестирование разработанной программной системы.

1. Анализ предметной области

1.1. Описание предметной области

Данная работа основывалась на использовании методов машинного обучения, а также компьютерного зрения с использованием искусственных нейронных сетей. Применение такой технологии можно найти в системах контроля доступа в помещениях, контроля доступа, подсчета посещаемости,

проверки безопасности на вокзалах и т.д. Более конкретная сфера применения данной технологии — обнаружение людей, которые не соблюдают карантинные меры, социальное дистанцирование и не носят защитные медицинские маски, а также последующего запрета им доступа в частные или общественные места.

С точки зрения актуальности исследования, в рамках социальной направленности, результаты позволяют в значительной степени упростить и ускорить процесс контроля доступа или допуска на определенную территорию или общественное место таких как: метро, аэропорт, университет и т.д.

1.2. Машинное обучение

Машинное обучение - это набор алгоритмов и инструментов, которые помогают машинам понимать закономерности в данных и использовать эту базовую структуру для выполнения действий и принятия решения о поставленной задаче. Есть много способов, которыми машины стремятся понять закономерности, лежащие в основе данных.

Термин машинное обучение был придуман в 1959 году Артуром Самуэлем, американским сотрудником IBM и пионером в области компьютерных игр и искусственного интеллекта [4][7]. Репрезентативной книгой исследований машинного обучения в 1960-х годах была книга Нильссона об обучающих машинах, посвященная в основном машинному обучению для классификации шаблонов [10]. Интерес, связанный с распознаванием образов, сохранялся и в 1970-х годах, как описано Дудой и Хартом в 1973 году [4]. В 1981 году был представлен отчет об использовании обучающих стратегий, так что нейронная сеть учится распознавать 40 символов с компьютерного терминала [12].

1.3. Глубокое обучение

Глубокое обучение - это ветвь машинного обучения, полностью

основанная на искусственных нейронных сетях, поскольку нейронная сеть будет имитировать человеческий мозг, поэтому глубокое обучение также является своего рода имитацией человеческого мозга. В глубоком обучении нам не нужно все явно программировать. Концепция глубокого обучения не нова. Он существует уже пару лет. Сейчас это шумиха, потому что раньше у нас не было такой вычислительной мощности и большого количества данных. Как и в последние 20 лет, вычислительная мощность растет в геометрической прогрессии, поэтому на первый план вышли глубокое обучение и машинное обучение. Формальное определение глубокого обучения - нейроны.

Несмотря на то, что глубокое обучение получило широкое распространение только в недавнее время, попытки обучать алгоритмы, основанные на многослойных нейронных сетях, осуществлялись с 60-х годов 20 века [2]. В 1980 году Кунихико Фукушима представил модель Неокогнитрон, предназначенную для распознавания визуальных образов [3].

Искусственный интеллект, особенно глубокое обучение, представляет собой значительный и, вполне возможно, революционный скачок вперед в улучшении технологий, поддерживающих здравоохранение. Такие компьютерные алгоритмы представляют собой инструменты, которые можно обучить для улучшения качества ухода за пациентами за счет повышения точности диагностики и снижения нагрузки на медицинских работников. Они предлагают возможности для автоматизации и прогнозирования, ранее невиданные в здравоохранении. Для поиска корреляций и закономерностей в клинических данных можно использовать очень большие наборы данных, что, вероятно, невозможно при использовании человеческого мозга без посторонней помощи или традиционных подходов. Наконец, системы глубокого обучения могут помочь медицинским организациям более эффективно взаимодействовать с персоналом и пациентами.

1.4. Переобучение нейронных сетей и методы предотвращения переобучения

Глубокие нейронные сети нацелены на изучение и обобщение закономерностей, обнаруженных в обучающих данных, чтобы они могли одинаково работать с тестовыми или новыми данными. Хотя это идеальное состояние или сценарий, который можно ожидать от глубоких нейронных сетей, это легче сказать, чем сделать, особенно для таких сетей, как сверточные нейронные сети, рекуррентные нейронные сети и т. Д., Которые состоят из миллионов или миллиардов настраиваемых параметров, которые уязвимы для переобучения.

Переобучение происходит, когда модель обучается на скрытых случайных закономерностях в обучающих данных, но отсутствующих в генеральной совокупности. Переобучение характеризует обобщающую способность модели.

Для решения проблемы переобучения глубоких нейронных сетей применяют следующие техники регуляризации:

- 1) L1- и L2-регуляризация весов
- 2) Ранняя остановка обучения
- 3) Метод “dropout”

Смысл первой техники заключается в ограничение абсолютных значений весов нейронной сети. Для этого к функции потерь применяется один из регуляризаторов:

- 1) L1-регуляризатор – сумма квадратов весов $L1 = \lambda \sum \omega^2$
- 2) L2-регулятор – сумма модулей весов $L2 = \lambda \sum |\omega|$

Вторая техника является довольно простой. Необходимо отложить часть обучающей выборки в валидационную выборку. Затем на каждой эпохе обучения вычислять ошибку на этой выборке – если ошибка начинает расти после определенного минимума, значит, нейронная сеть начала переобучаться и нужно остановить процесс обучения. Есть данные, что данная техника в некоторой степени близка к L2-регуляризации [7]. Также анализ ошибки на валидационной выборке позволяет оценивать обобщающую способность модели.

Третий метод также нашел широкое применение для предупреждения переобучения нейронных сетей, т.к. является более эффективным, чем первый, и часто применяется в первую очередь. Идея метода “dropout” заключается в следующем – для каждого нейрона кроме нейронов выходного слоя определяется вероятность p . С этой вероятностью нейрон отключается. Алгоритм обучения становится таким: на каждом сэмпле x из обучающей выборки вычисляется, будет ли активен нейрон, и в зависимости от результата нейрон либо остается активным, либо отключается и его выход устанавливается равным нулю. Ноль на выходе приводит к тому, что нейрон отсутствует в вычислительном графе. Прямой проход, и обратное распространение градиента не идут дальше неактивного нейрона [8]. Эксперименты [9], и практика обучения нейронных сетей показывают, что метод “dropout” действительно дает очень серьезные улучшения в качестве обученной модели в самых разных приложениях.

1.5. Свёрточные нейронные сети

Сверточная нейронная сеть (также CNN или ConvNet) является одним из наиболее популярных алгоритмов в глубоком обучении, это такой вид машинного обучения, при котором модель учится выполнять задачи классификации непосредственно на изображении, видео, тексте или звуке.

Сверточные нейронные сети особенно полезны для поиска шаблонов (паттернов) на изображениях для распознавания объектов, лиц и сцен. Они учатся непосредственно на изображениях, используя шаблоны для классификации изображений и устраняя потребность в ручном извлечении признаков.

Приложения, требующие распознавания объектов и компьютерного зрения, такие как беспилотные автомобили и приложения для распознавания лиц, во многом зависят от CNN. В зависимости от вашего приложения вы можете построить свою CNN с нуля или использовать готовую модель на ваших данных.

Главная идея свёрточных нейронных сетей заключается в том, обработка части изображения должна происходить независимо от местоположения этой части. Реализуется эта идея явно – мы покрываем входное изображение небольшими окнами (например, 5x5 пикселей) и выделяем признаки в каждом. В таком окне небольшой нейронной сетью. При этом признаки выделяются в каждом окне одни и те же, значит, будет одна нейронная сеть. У нее будет $5 \times 5 = 25$ входов. Из каждого изображения получится много разных входов для этой нейронной сети. После этого выход можно будет снова представить в виде изображения, заменив окна 5×5 на их центральные пиксели. На этом изображении можно будет применить второй свёрточный слой с уже другой нейронной сетью. В каждом свёрточном слое будет немного свободных параметров по сравнению с полносвязными нейронными сетями.

Обычно исходное изображение представляет собой три матрицы (по цветовым каналам RGB) – красный, зеленый и голубой. Такие матрицы называют каналами. Такие же матрицы будут получаться после каждого свёрточного слоя. В них присутствует пространственная структура, соответствующая исходной картинке. Каналов становится больше. Значения

каждого признака, которые мы выделили из окон в исходном изображении, теперь будут представлять собой целую матрицу. Такие матрицы принято называть картами признаков.

Операция свёртки — это линейное преобразование входных данных особого вида. Пусть x^l — карта признаков в слое l . Тогда результат двумерной свертки с ядром размера $2d+1$ и матрицей весов W размера $(2d+1) \times (2d+1)$ на следующем слое будет равен:

$$y_{i,j}^l = \sum_{-d \leq a, b \leq d} W_{a,b} x_{i+a, j+b}^l \quad (1)$$

где $y_{i,j}^l$ — результат свертки на уровне l , а $x_{i,j}^l$ — вход свёртки, он же выход предыдущего слоя. Для получения результата на следующем слое, мы применяем свёртку к квадратному окну с предыдущего слоя, то есть скалярно умножаем пиксели окна на вектор свёртки.

Операция свёртки обладает следующими свойствами:

- 1) Операция свертки обладает свойством разреженности, так как значение каждого нейрона очередного слоя зависит только от небольшой доли входных нейронов;
- 2) Свёртка многократно переиспользует одни и те же веса, так как они повторно применяются к различным участкам входа.
- 3) Свёртка сохраняет структуру входного изображения, так как применяется к каждому участку входных данных в отдельности;

После операции свёртки можно, применить ту или иную нелинейную функцию h : она будет просто применяться к каждому элементу полученного тензора по отдельности. В классическом сверточном слое, кроме линейной свертки и следующей за ней функцией активации, есть и еще одна операция:

ПУЛИНГ.

В свёрточных сетях важнее наличие или отсутствие какого-либо признака, чем его точные координаты. Например, при распознавании лиц важно понять, есть ли на изображении лицо и кому оно принадлежит, а не узнать, где оно расположено на изображении. Поэтому мы можем позволить себе обобщить выделяемые признаки, при этом потеряв часть информации об их расположении. К тому же это дает нам возможность сократить размерность [6].

В основном в качестве операции пулинга применяется операция взятия максимума (max-pooling). Формула для расчёта пулинга:

$$x_{i,j}^{l+1} = \max_{-d \leq a \leq d, -d \leq b \leq d} z_{i+a,j+b}^l \quad (2)$$

Пример работы операции макс-пулинг представлен на рисунке 1 (см. ниже).

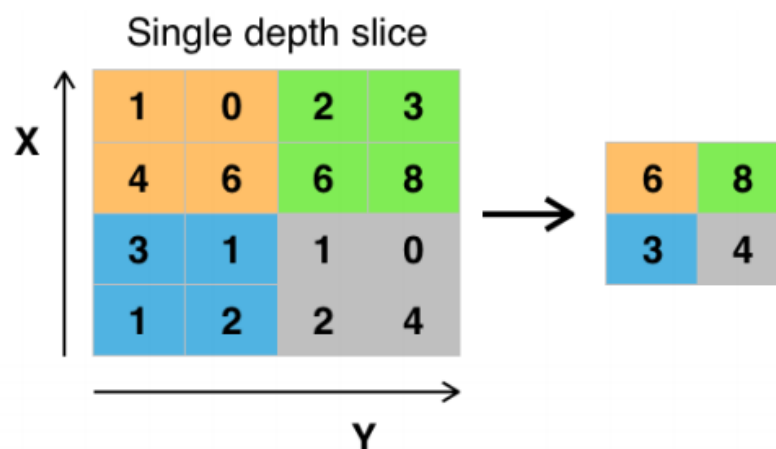


Рисунок 1 – Макс-пулинг

Другим важным элементом свёрточных нейронных сетей является функция активации, о которой было сказано ранее. Функция активации после слоя свёртки является источником нелинейности, которая дает нам возможность обучать нейронные сети. Функция активации должна отвечать следующим критериям:

- Быть монотонно неубывающей
- Быть ограниченной, т.е. стремиться к нулю при $x \rightarrow -\infty$
- Быть дифференцируемой

Существует множество функция активации: логистическая сигмоида, гиперболический тангенс и другие. Но наиболее популярной функцией активации для свёрточных нейронных сетей является rectified linear units (ReLU) [10]. Формула ReLU имеет следующий вид:

$$ReLU(x) = \begin{cases} 0, & \text{если } x < 0, \\ x, & \text{если } x \geq 0. \end{cases}$$

Функция ReLU более эффективна, чем логистическая регрессия или гиперболический тангенс, т.к. вычисление её производной не требует сложных расчётов: если $x < 0$, тогда производная равна нулю, если $x \geq 0$, тогда производная равна единице. Это также позволяет уменьшить расходы на вычислительные ресурсы. Кроме классической реализации ReLU существуют её модификации, позволяющие получить лучший результат в некоторых задачах [6].

1.6. Системы распознавания лиц

Распознавание лиц - это программный алгоритм, используемый для проверки или идентификации личности человека путем обработки видеокadra или цифрового изображения, на котором видно лицо человека. Технологии распознавания лиц используют несколько различных методов, но обычно они сравнивают черты лица на изображении с лицами, содержащимися в базе данных.

Процессом распознавания лиц – это набор различных задач, которые

служат для идентификации лица по цифровому изображению или видео. В общем виде этот процесс происходит следующим образом: после того, как система получила изображение с камеры, первым шагом в системе является обнаружение лиц, в котором определяется, содержит ли изображение или видео какие-либо лица. Обнаружение лица также определяет расположение ключевых функций (так называемых «ориентирами»), таких как глаза, нос и т. д. Эти ориентиры также могут использоваться для определения местоположения лица на изображении или видео и определения его размера и ориентации. Далее с помощью алгоритмов определяются границы лица (этап обнаружения). Соответственно следует этап распознавания, на котором лицо трансформируется (изменяется его яркость, оно выравнивается, масштабируется, и т.п.) и приводится к некоторому заданному виду. В заключительном этапе происходит обнаружение признаков и непосредственно сравнение их с признаками из баз данных. Данный этап сравнения называется идентификация или верификация.

Верификация: этот процесс сопоставления образцов по схеме «один к одному» («1:1»). Для определения личности система сравнивает биометрический образец с одним биометрическим шаблоном, хранящимся в базе данных, чтобы предсказать, являются ли они одинаковыми».

Идентификация: данный процесс сопоставления по схеме «один ко многим» (1: N) – где система сравнивает лицо на изображении и биометрический образец со всеми шаблонами лиц, хранящимися в базе данных.

На рисунке 2 (см. ниже) изображен общий алгоритм распознавания лиц по изображению или видео.



Рисунок 2 – Алгоритм распознавания лиц

1.7. Методы обнаружения лиц

Существует множество алгоритмов обнаружения лиц. Большинство из них доступны в виде статей с описанием математических моделей, а другие представлены в виде программной реализации. Изображение в виде цифровых данных с камер передается на компьютер – оно обрабатывается с помощью специального алгоритма, который определяет расположение области лица по его основным чертам (глазам, рту, бровям, носу и т. д.). Подобных методов обнаружения лиц существует много и большинство из них представляют собой комбинацию других методов.

В данном разделе приведены классификация алгоритмов обнаружения лиц, а также описаны преимущества и недостатки каждой группы алгоритмов. Но все их можно разбить на две категории: методы на основе знаний и методы обнаружения лица по внешним признакам. Классификация методов обнаружения приведена на рисунке 3 (см. ниже).

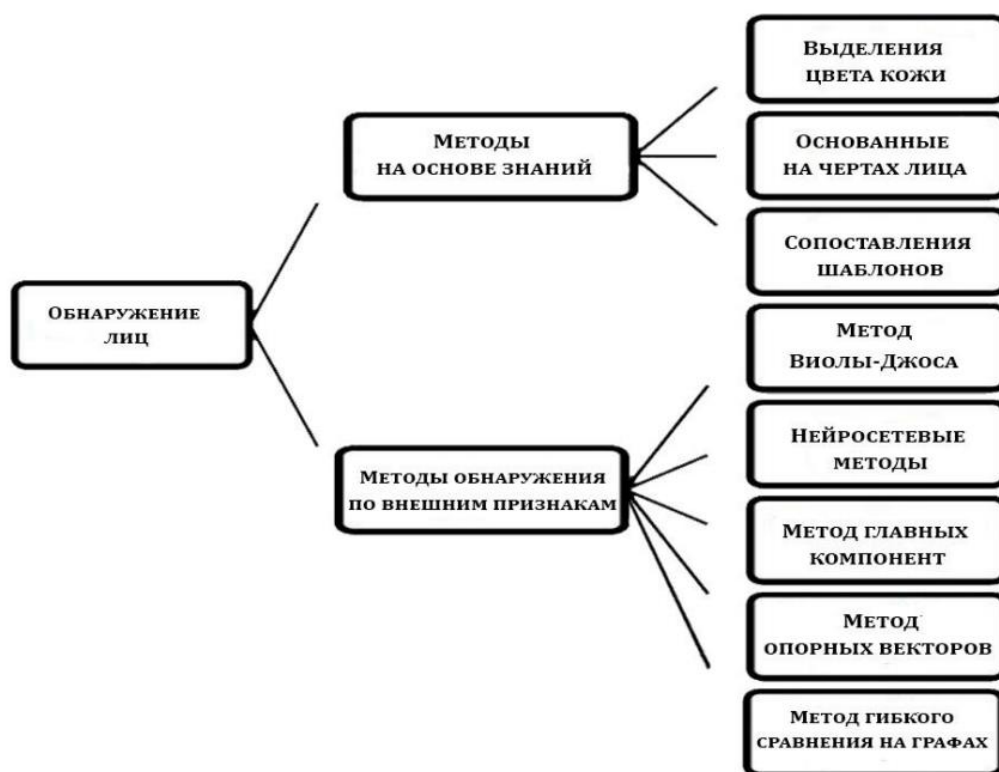


Рисунок 3 – Классификация методов обнаружения лиц

Основанные на знаниях методы зависят от набора правил, и он основан на человеческих знаниях для обнаружения лиц. Эти методы используют информацию о лице (лицо должно иметь нос, глаза и рот на определенном расстоянии друг от друга), форме, цвете кожи, текстуре или его чертах. Все правила — это формализованные знания, которыми руководствуется человек. Например, основные правила: области глаз, носа и рта отличаются по яркости относительно остальной части лица; глаза на лице всегда располагаются симметрично относительно друг друга. Большая проблема с этими методами - сложность в создании соответствующего набора правил. Там может быть много ложных срабатываний, если правила были слишком общими или слишком подробными. Одного этого подхода недостаточно и невозможно найти много лиц на нескольких изображениях.

К этой же группе методов относят более общий метод - метод шаблон соответствия. Метод сопоставления с шаблоном использует предварительно

определенные или параметризованные шаблоны лица для определения или обнаружения лиц по корреляции между шаблонами и входными изображениями. Например, человеческое лицо можно разделить на глаза, контур лица, нос и рот. В данном методе, описания свойств отдельных областей лица и их заданному взаимному расположению, которые поможет по шаблонам и стандарт лица найти соответствия с исходным изображением.

Методы на основе знаний получили довольно широкое распространение и имеют неплохие показатели, однако они показывают хорошие результаты только на изображениях с хорошим разрешением, без шумов и с несложным фоном. На кадрах с видеопотока или камер, установленных в общественных местах, где возможны различные ракурсы и повороты лиц, а также меняющееся освещение и большое количество мешающих объектов на заднем плане, существует большая вероятность возникновения ошибок.

Метод обнаружения лиц основанный по внешним признакам, зависит от набора обучаемых изображений лица делегата для определения моделей лица. Подход, основанный на внешности, лучше, чем другие способы исполнения. В общем, метод, основанный по внешним признакам, основывается на методах статистического анализа и машинного обучения, чтобы найти соответствующие характеристики изображений лица. Этот метод также используется для выделения признаков и распознавания лиц. Соответственно методы данной группы не имеют выше отмеченных недостатков и их стали чаще применять в системах видеонаблюдения. Обнаружение лиц в таких методах осуществляется перебором всех фрагментов изображения для определения, к какому классу относится изображение: к классу изображений без лица или к классу содержащих лицо.

Из-за большого объема работы, методы обладают избыточностью и большой вычислительной сложностью. Для того чтобы уменьшить количество

вычислений и ускорить процесс отыскания лиц, применяются дополнительные методы для сокращения количества рассматриваемых фрагментов.

Наиболее актуальных методов обнаружения лиц рассмотрены ниже:

1.8. Метод Виолы-Джоса

Метод Виолы-Джоса (Viola–Jones object detection). Метод был назван в честь двух исследователей компьютерного зрения, предложивших этот метод в 2001 году, Паула Виолы и Майкла Джонса в своей статье «Быстрое обнаружение объектов с использованием усиленного каскада простых функций». Несмотря на то, что метод Viola-Jones устарел, он довольно мощный, и его применение оказалось исключительно заметным при обнаружении лиц в реальном времени. Этот алгоритм может обнаруживать лица в реальном времени с впечатляющей скоростью. [8]. У данного метода есть множество реализаций, в том числе в составе библиотеки компьютерного зрения OpenCV[9].

Основные этапы работы метода:

- используются изображения в интегральном представлении, что позволяет вычислять быстро необходимые объекты;
- используются признаки Хаара, с помощью которых происходит поиск нужного объекта (в данном контексте, лица и его черт);
- используется бустинг (от англ. boost – улучшение, усиление) для выбора наиболее подходящих признаков для искомого объекта на данной части изображения; все признаки поступают на вход классификатора, который даёт результат «верно» либо «ложь»;
- используются каскады признаков для быстрого отбрасывания окон, где не найдено лицо.

Элементы каскада Хаара состоит из темных и светлых областей. Он производит одно значение, взяв сумму интенсивностей светлых областей и вычитая ее из суммы интенсивностей темных областей. Эти признаки Хаара дают точечное значение перепада яркости по оси X и Y соответственно. Поэтому общий признак Хаара для распознавания лиц представляет набор двух смежных прямоугольников, которые лежат выше глаз и на щеках. Значение признака вычисляется по формуле:

$$F = X - Y \quad (1)$$

где X – это сумма значений яркостей точек закрываемых светлой частью признака, а Y – это сумма значений яркостей точек закрываемых темной частью признака.

Интегральное представление изображений можно представить в виде матрицы, размеры которой совпадают с размерами исходного изображения I, и, каждом элементе ее хранится сумма интенсивностей всех пикселей, находящихся левее и выше данного элемента. Элементы матрицы рассчитываются по следующей формуле:

$$L(x, y) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(i, j) \quad (2)$$

где $I(i, j)$ — яркость пикселя исходного изображения.

Каждый элемент матрицы $L(x, y)$ представляет собой сумму пикселей в прямоугольнике от (0,0) до (x,y), т.е. значение каждого пикселя (x,y) равно сумме значений всех пикселей левее и выше данного пикселя (x,y).

Для того, чтобы вычислить сумму прямоугольной области в интегральном представлении изображения требуется всего 4 операции обращения к массиву и 3 арифметические операции. Это позволяет быстро рассчитывать признаки Хаара для изображения в обучении и распознавании [7].

По такой интегральной матрице можно очень быстро вычислить сумму пикселей произвольного прямоугольника, произвольной площади. Пусть в прямоугольнике ABCD есть интересующий нас объект D:

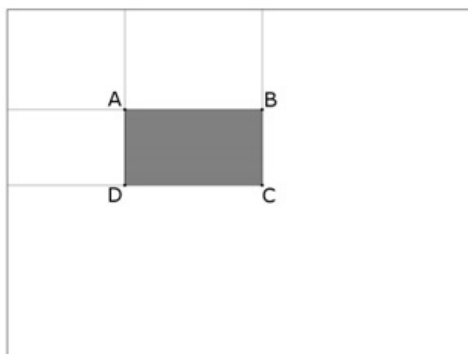


Рисунок 4 – прямоугольную область с использованием интегрального изображения.

Сумму внутри прямоугольника ABCD можно выразить через суммы и разности смежных прямоугольников по формуле:

$$S(ABCD) = L(A) + L(C) - L(B) - L(D) \quad (3)$$

Алгоритм AdaBoost (Adaptive Boosting) – это алгоритм машинного обучения для выбора лучшего подмножества функций среди всех доступных функций. Результатом работы алгоритма является классификатор (он же функция прогнозирования, функция гипотез), называемый «сильным классификатором» [7]. Сильный классификатор состоит из линейных комбинаций «Слабых классификаторов» (лучших характеристик). С высокого уровня, чтобы найти эти слабые классификаторы, алгоритм запускается для T итераций, где T - количество слабых классификаторов, которые нужно найти. На каждой итерации алгоритм находит частоту ошибок для всех функций, а затем выбирает функцию с самой низкой частотой ошибок для этой итерации. AdaBoost выбирает набор слабых классификаторов для объединения и присваивает каждому из них свой вес. Эта взвешенная комбинация и есть сильным классификатором.

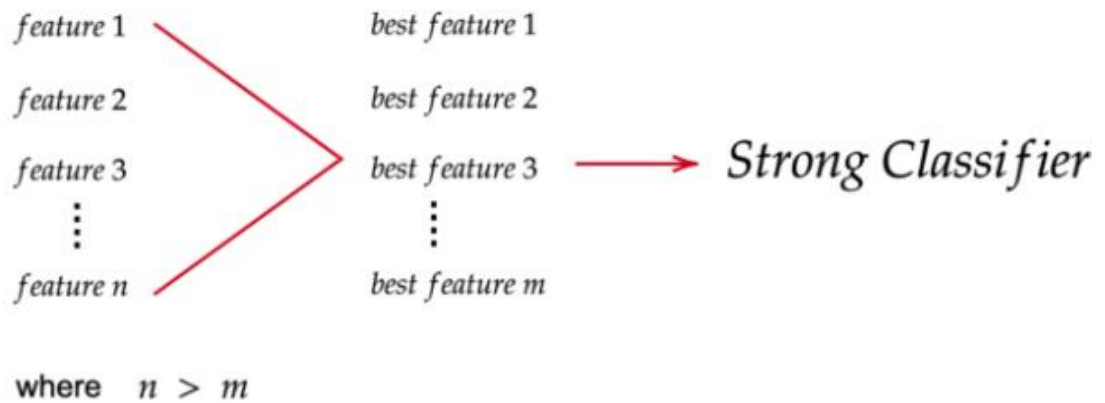


Рисунок 5: Цель использования алгоритма AdaBoost

Для каждого признака, слабый классификатор определяет оптимальную пороговую функцию классификации, обеспечивающую минимальное число неправильно классифицированных примеров. Слабый классификатор h таким образом, состоит из значений признака объекта f , порога и паритета p , с указанием направления знака неравенства.

В результате работы алгоритма бустинга на каждой итерации формируется простой классификатор вида:

$$h_j(z) = \begin{cases} 1, & \text{если } p_j f_j(z) < p_j Q_j \\ 0, & \text{иначе} \end{cases} \quad (4)$$

Где, p_j показывает направление знака неравенства, Q_j – пороговое значение по j , $f_j(z)$ – вычисленное значение признака, z – окно изображения.

Каскадный классификатор – это многоступенчатый классификатор, который может выполнять обнаружение быстро и точно. Каждый этап состоит из сильного классификатора, созданного алгоритмом AdaBoost. От этапа к этапу количество слабых классификаторов в сильном классификаторе увеличивается [7]. Входные данные оцениваются последовательно (поэтапно). Если классификатор для определенного этапа дает отрицательный результат, ввод немедленно отбрасывается. Если выход положительный, вход перенаправляется на следующий этап, как приведено на рисунке 6 (см ниже).

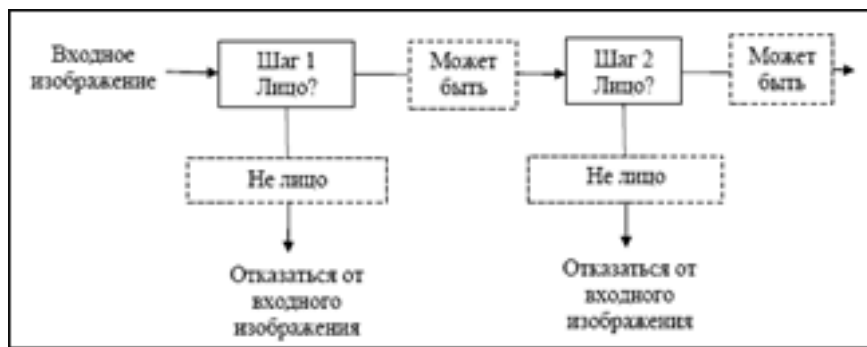


Рисунок 6: Каскадный классификатор.

Когда фильтр пропускает область изображения, она переходит к следующему фильтру в последовательности. Когда вся область изображения, прошли через все фильтры, потом классифицируются как «лицо», а остальные классифицируются как «не лицо».

Преимущества метода Виолы-Джонса:

- Обладает высокой точностью обнаружения повернутых лиц под разными углами (до 60-65) градусов;
- Быстродействие работы;

Недостатки:

- Большое количество близко расположенных друг к другу результатов из-за применения различных масштабов и скользящего окна;
- Длительное время обучения. Алгоритму необходимо проанализировать большое количество тестовых изображений.

Машины опорных векторов - это линейные классификаторы, которые максимизируют разницу между гиперплоскостью решения и примерами в обучающем наборе. Суть метода заключается в сопоставлении графов, которые

описывают лица (лицо представляется в виде сетки с индивидуальным расположением вершин и ребер).

Процедура распознавания происходит следующим образом - граф, характеризующий параметр распознавания, остается без изменения, а другие деформируются под влиянием структура лица с основными антропометрическими точками (расстояние между ушами, глазами, ширина губ, линия носа и т.п.) [10]. Чем больше этих точек используется, соответственно тем точнее будет результат распознавания.

Недостатки метода:

- Сложность алгоритма распознавания приводит к необходимости использования значительных вычислительных мощностей;
- Сложная процедура введения новых шаблонов в базу данных;
- Быстродействие аналитической системы обратно пропорционально размерам баз данных.

Скрытые Марковские модели (СММ) – это набор статистических моделей, используемых для характеристики статистических свойств сигнала и учитывают их пространственные характеристики для сравнения объектов с шаблонами из базы [10].

Скрытые Марковские модели особенно известны своим применением в распознавании одномерных образов, таких как распознавание речи, анализ музыкальной партитуры и задачи определения последовательности в биоинформатике. Совсем недавно они были применены к более сложным 2D-задачам, и в этом обзоре основное внимание уделяется их использованию в области *автоматическое распознавание лиц*, отслеживая эволюцию использования ТММ с начала 1990-х годов до наших дней [10]. Во время распознавания человека, происходит проверка все сгенерированные

Марковские модели и ищется из наблюдаемых вероятность того, что последовательность наблюдений для объекта сгенерирована соответствующей моделью.

Недостатки:

- Не большая скорость работы;
- Система может оптимизировать время обработки данных и отклика, но не может минимизировать время перебора других моделей;
- Низкая различающая способность и не оптимальный алгоритм обучения.

Метод главных компонент (PCA). Алгоритмы, такие как PCA и дискриминант Фишера, могут использоваться для определения подпространства, представляющего шаблоны лица. Главной целью PCA является уменьшение пространства признаков без значимой потери информации и так, чтобы оно как можно лучше описывало «типичные» образы, принадлежащие множеству лиц. Существует обученный классификатор, который идентифицирует экземпляры класса шаблонов из шаблонов фонового изображения. Алгоритма используют для того, чтобы представить лицо как вектор малой размерности, который затем сравнивается с эталонными векторами из представленных в базе.

Нейросетевые методы. Нейронные сети успешно столкнулись со многими проблемами обнаружения, такими как обнаружение объектов, обнаружение лиц, обнаружение эмоций, распознавание лиц и т. Д. Довольно распространенные методы, которые включают в себя около десятка различных алгоритмов. может использоваться для классификации изображений лиц на основе таких характеристик, как расстояния между некоторыми

специфическими частями лица (нос, рот, глаза). Также существуют гибридные системы, например, объединение с марковской моделью. В классической МНС межслойные нейронные соединения полносвязны, и изображение представлено в виде одномерного вектора, хотя оно двумерно. Архитектура сверточной НС направлена на преодоление этих недостатков [7]. В ней используются локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов), общие веса (обеспечивают детектирование некоторых черт в любом месте изображения) и иерархическая организация с пространственными подвыборками (spatial subsampling). Помимо этого, любая обученная МНС в некоторой мере может определять принадлежность образов к "своим" классам, ее можно специально обучить надежному детектированию определенных классов [10]. В этом случае выходными классами будут классы, принадлежащие и не принадлежащие к заданному типу образов.

Нейросетевые методы достигают одни из самых лучших результатов для распознавания, но считаются наиболее сложными для реализации.

Преимущества данного метода:

— Высокая точность обнаружения при правильной настройке параметров сети.

Недостатки:

— Сложная процедура внесения изменений в процессе обучения (любое изменение требует переобучения сети);

— Трудность формирования архитектуры сети;

— Требуется высокая вычислительная мощность.

1.9. Методы предварительной обработки изображений

Для повышения эффективности распознавания и качества выделения

лиц, в системах распознавания проводят этап предобработки входных изображений.

После того, как система находит лицо в кадре и определяет положение головы, размер, позу и его ключевые черты, оно, как правило, подвергается нормализации: кодируется, масштабируется, преобразуется до горизонтального положения линии, соединяющей центры глаз. Также, при предобработке лица, применяются разнообразные фильтры для снижения уровня шума (медианный, гауссовский и пр.).

Методы предварительной обработки довольно разнообразны и зависят от задач исследований. Наиболее часто встречаются из них:

Нелинейные фильтры – фильтры, которые используют для удаления импульсного шума (отдельных точек с максимальной (белой) или минимальной (черной) яркостью) на изображении. Действия таких фильтров заключается в определении позиции каждого импульса и замене их значениями фиксированной или случайной величины.

Фильтр Гаусса – фильтр размытия изображения, который используют, если на изображении существуют мелкие детали, которые не требуют отделения от фона и их можно размыть.

Медианные фильтры – фильтры, использующиеся для сохранения перепадов яркости контуров и подавления импульсных шумов. Медианные фильтры достаточно хорошо работают в таких случаях, при которых плотность шума невелика.

1.10. Методы распознавания лиц

Другая важная часть систем автоматического распознавания лиц — это непосредственно сами алгоритмы распознавания человека. На сегодняшний день таких алгоритмов много, и каждый из них имеет свою специфику, свою

скорость работы и свою надежность распознавания. Алгоритмы распознавания делятся на две категории, в зависимости от применяемой технологии распознавания – двумерные, в которых распознавание происходит по геометрии лица (2D-технологии) и трехмерные, в которых распознавание происходит по строению черепа (3D-технологии).

Системы 2D-распознавания работают с «плоскими», двухмерными изображениями и распознают лицо, анализируя его текстуру и участки лица с высокой контрастностью, поэтому при нарушении освещения или положения лица его распознавание сильно затрудняется.

Системы 3D-распознавания же более устойчивы к таким изменениям, т. к. при создании модели лица, в них учитываются особенности строения черепа. Но тем не менее, хоть системы 3D и имеют такой большой плюс, из-за отсутствия возможности обслуживать большое количество пользователей в режиме идентификации и из-за невысокой скорости, 3D-технологии пока не получили широкого применения. Также, 3D-технологии распознавания требуют больших вычислительных ресурсов, и стоимость оборудования для таких систем намного выше, чем у 2D [5].

2. Разработка системы распознавания лиц в защитной маске

2.1. OpenCV

OpenCV (англ. Open Source Computer Vision Library, библиотека компьютерного зрения с открытым исходным кодом) – библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения с открытым кодом, поддерживающая множество платформ, включая Windows, MacOS и Linux. Также эта библиотека существует и для многих языков программирования. Но наиболее часто она используется для написания приложений машинного обучения на языке Python, особенно в сфере компьютерного зрения. Реализована на C/C++, также разрабатывается для

Python, Java, Ruby, Matlab, Lua и других языков (рисунок 7).

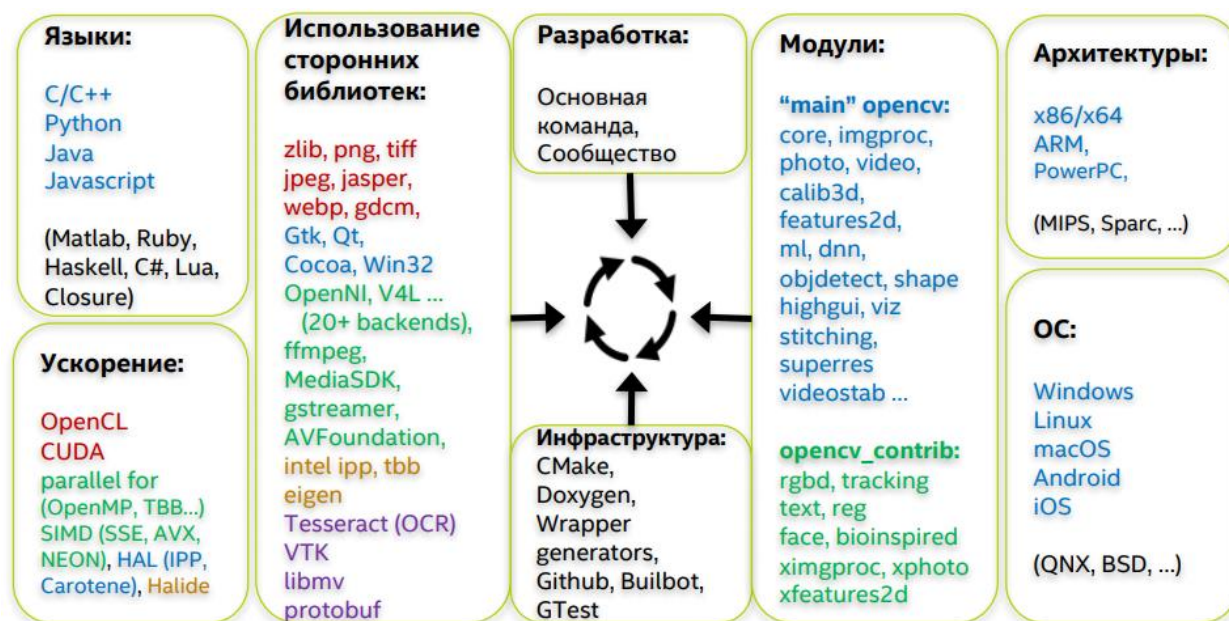


Рисунок 7 – Карта применение OpenCV.

В данной работе OpenCV исполняет роль базовой компоненты, предоставляющей инструментарий для работы с входными данными, представленными видеопоследовательностью. Данная библиотека позволяет представить видеоряд в качестве набора кадров (фреймов), каждый из которых предполагается рассматривать как отдельное изображение.

Также OpenCV поддерживает работу с потоковым видео, полученным напрямую с камеры, что открывает для разрабатываемой системы потенциал обработки и анализа входных данных в режиме реального времени, что является крайне важным аспектом в проектировании системы компьютерного зрения в области транспортных средств. Может свободно использоваться в академических и коммерческих целях. Также, данная библиотека содержит удобный функционал по работе с изображениями, алгоритмы их обработки, распознавания лиц и пр. Именно поэтому было принято решение об использовании OpenCV [9].

2.2.Dlib

Dlib – это современная библиотека, представляющая собой набор методов и алгоритмов машинного обучения, а также инструментарий для создания сложного программного обеспечения для решения задач в областях промышленности, робототехники, мобильных телефонов, а также высокопроизводительных вычислительных средах

В данной работе используется метод «simple_object_detector», который позволяет создать и обучить модель для детектирования определённых объектов на изображении. Для тренировки такой модели в качестве обучающей выборки на вход подаются предварительно размеченные изображения, содержащие детектируемый объект. Далее такие изображения анализируются с помощью детектора гистограммы направленных градиентов (рис. 9), позволяющего рассчитать градиент в каждой точке пространства, и в зависимости от направления этого градиента записать в гистограмму некоторое значение.



Входное изображение

Гистограмма направленных градиентов

Рисунок 8 Гистограмма направленных градиентов (HOG).

С помощью такого анализа входных данных формируется набор векторов, каждый из которых описывает точку в n -мерном пространстве.

2.3.TensorFlow

TensorFlow – открытая программная библиотека для машинного обучения, разработанная компанией Google для решения задач построения и тренировки нейронной сети с целью автоматического нахождения и классификации образов. Основной API был реализован для Python. Этот фреймворк позволяет абстрагироваться от внутренних деталей машинного обучения и сосредоточиться непосредственно на решении задачи, позволяет создавать, обучать и использовать нейронные сети любого известного типа [18].

Данная библиотека использует реализацию ряда узлов обработки, представляющих собой математическую операцию. Ряд таких узлов называется «графом» (рисунок 9).

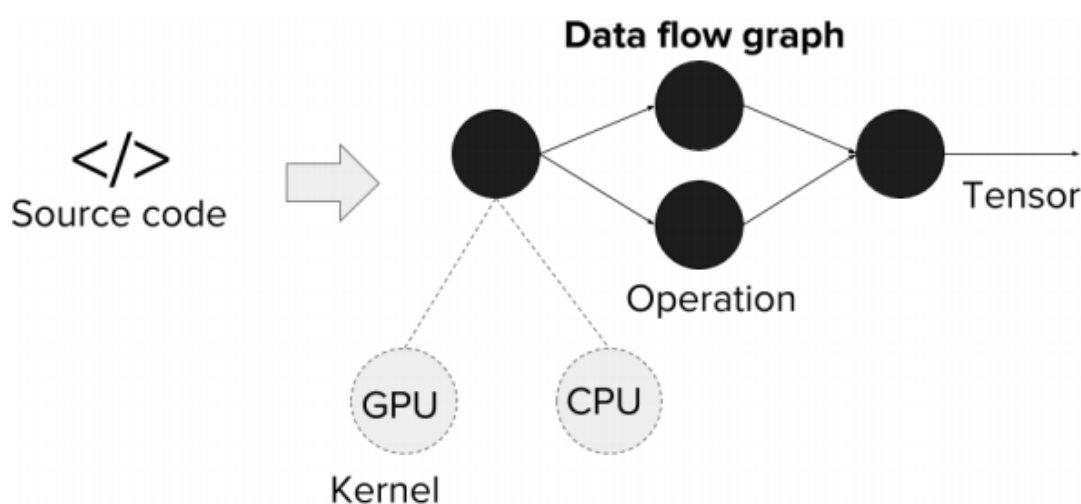


Рисунок 9– Схема работы TensorFlow.

2.4.NumPy

NumPy или Numerical Python — это библиотека с открытым исходным кодом для языка программирования Python, которая предоставляет общие математические и числовые операции в виде пре-скомпилированных, быстрых функций., которая предлагает следующее:

- Мощный N-мерный массив.

- Высокоуровневые функции.
- Инструменты для интеграции кода C/C++ и Fortran.
- Использование линейной алгебры, Преобразований Фурье и возможностей случайных чисел.

Она также предлагает эффективный многомерный контейнер общих данных. NumPy обеспечивает функциональность, которую можно сравнить с функционалом MatLab. NumPy предоставляет базовые методы для манипуляции с большими массивами и матрицами.

2.5. Используемые архитектуры нейронных сетей

Традиционный подход во всех алгоритмах обнаружения объектов заключается в использовании подхода скользящего окна, но это требует больших вычислительных ресурсов. YOLO преодолевает это, используя концепции ячеек сетки.

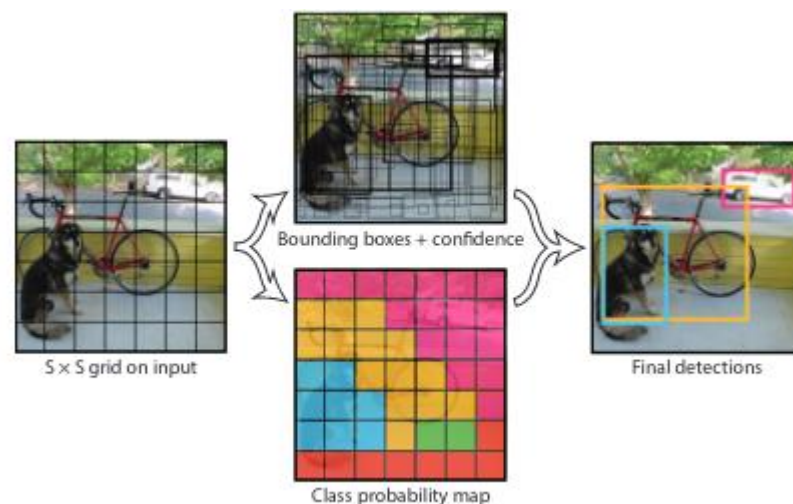


Рисунок 10 – Обнаружение объектов архитектуры YOLO

- YOLO использует ячейки сетки $S \times S$ (здесь $S = 7$). Если центр объекта попадает в ячейку сетки, эта ячейка сетки отвечает за обнаружение объекта.

- Каждая ячейка сетки дает B ($B = 2$) ограничивающих прямоугольников и оценку достоверности / оценку объектности этих прямоугольников. Оценка уверенности означает, насколько уверенно прямоугольник содержит объект.
- Каждый прогноз ограничивающей рамки состоит из координат x , y , w , h и оценки достоверности. Координаты (x, y) указывают центр прямоугольника относительно границ ячейки сетки. Координаты (w, h) указывают ширину и высоту относительно изображения (рисунок 11).

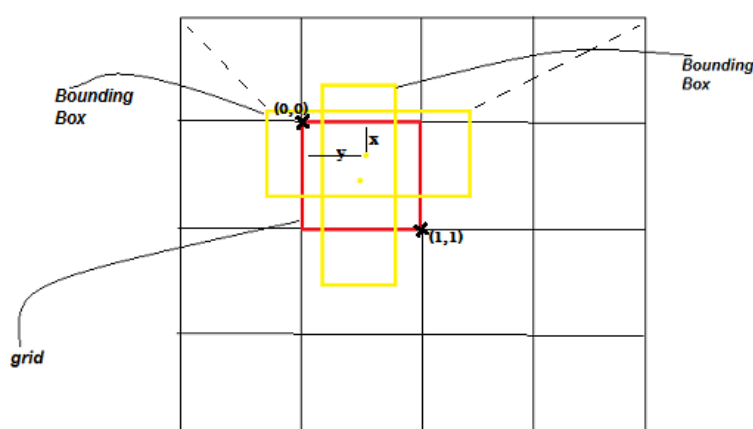


Рисунок 11 – Сверка изображений в YOLO

Сетевая архитектура YOLO вдохновлена архитектурой GoogLeNet. Сеть состоит из 24 сверточных слоев, за которыми следуют 2 полностью связанных слоя. Вместо начальных модулей, используемых GoogLeNet, мы просто используем уменьшающие слои 1×1 , за которыми следуют сверточные слои 3×3 .

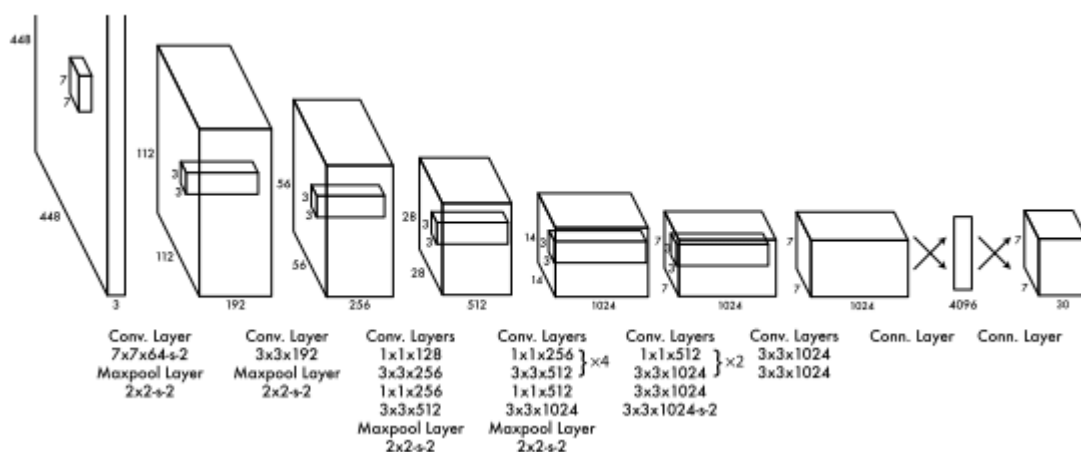


Рисунок 12 – Архитектура YOLO

2.6. Используемые наборы данных

MaskTheFace - это основанный на компьютерном зрении скрипт для маскировки лиц на изображениях. Он использует *dlib* для определения наклона лица и шести основных черт лица, необходимых для нанесения маски. На основании наклона лица из библиотеки маски выбирается соответствующий шаблон маски. Затем шаблонная маска трансформируется на основе шести ключевых особенностей, чтобы она идеально соответствовала лицу. Полную блок-схему можно увидеть ниже (рисунок 13).

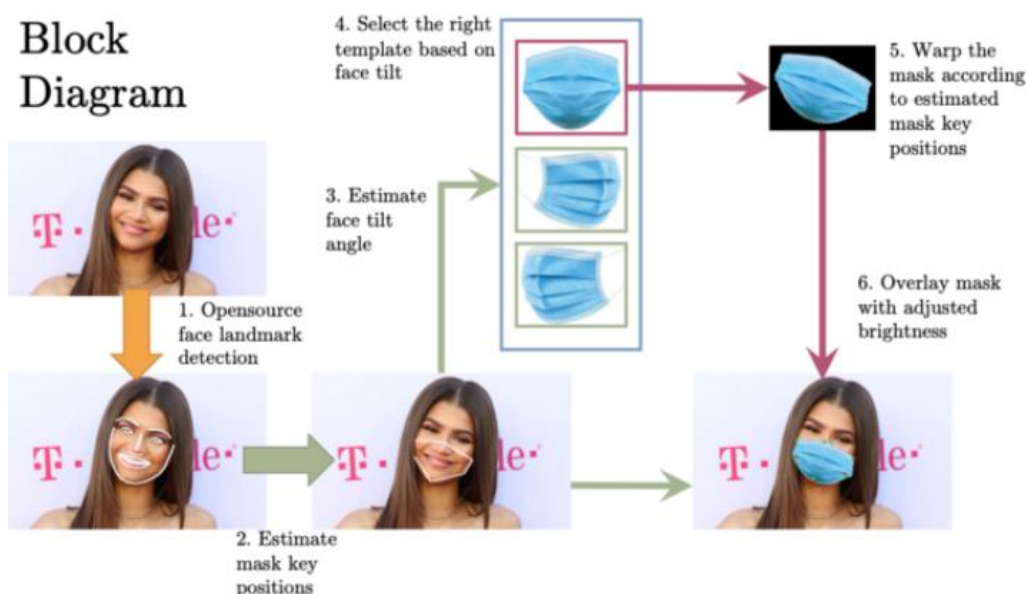


Рисунок 13 – схема создание датасета MaskTheFace

В настоящее время датасет MaskTheFace содержит следующие 4 вида масок (рисунок 14).



Рисунок 14 – Виды масок присутствующих в датасете.

MaskTheFace можно использовать для создания набора данных синтетической маски лица, который затем можно использовать для обучения различных прикладных сетей машинного обучения, таких как распознавание лиц для системы посещаемости в помещении, обнаружение масок, классификация масок и т. Д.

MFR2 - это небольшой набор данных с 53 личностями знаменитостей и политиков, в общей сложности 269 изображений, собранных из Интернета. Каждая личность имеет в среднем 5 изображений. Набор данных содержит как замаскированные, так и немаскированные лица лиц. Набор данных обрабатывается с точки зрения выравнивания лица и размеров изображения. Каждое изображение имеет размер (160x160x3). Примеры изображений из набора данных MFR2 и распределение масок можно увидеть ниже (рисунок 15).

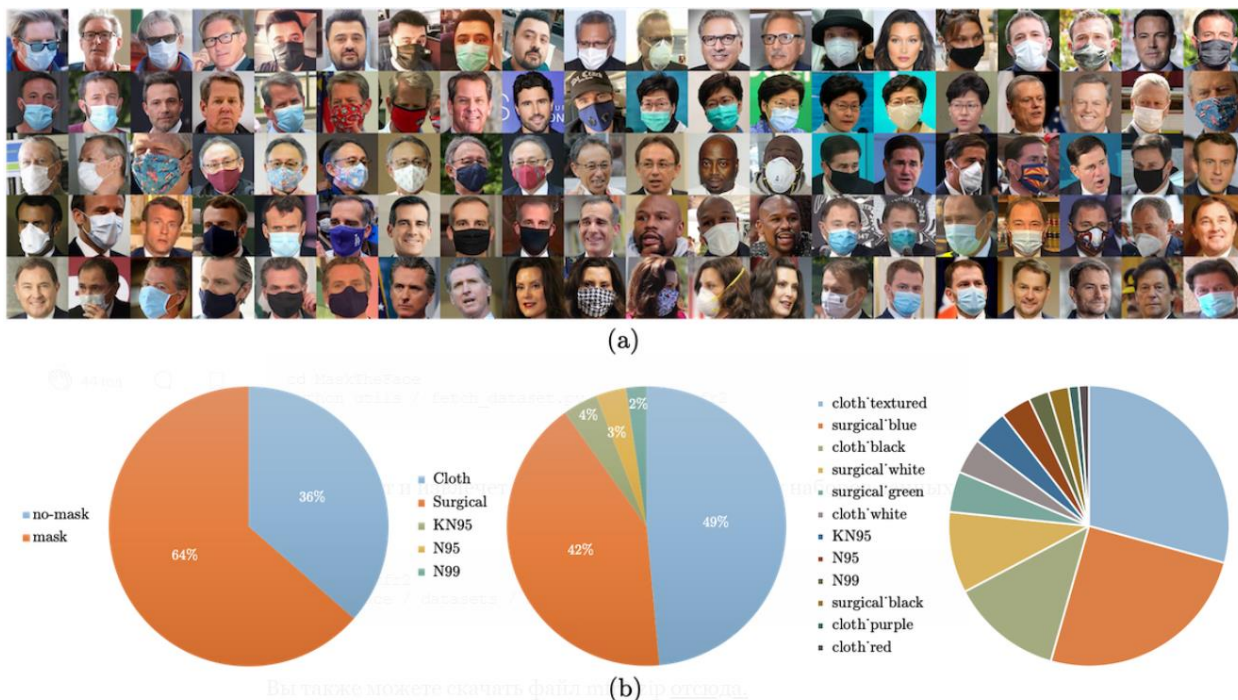


Рисунок 15 – Наборы данных MFR2.

2.7. Обработки кадров видеопотока

Как было отмечено ранее, обработка кадров видеопотока состоит из двух основных этапов. На первом этапе происходит обнаружение лиц в кадре, а на втором – обнаружение защитной маске на лице.

При разработке системы, в данной работе был использован метод обнаружения лиц по признакам каскада Хаара, а для распознавания защитной маски на лице был применен нейросетевой метод. В качестве датасета для обучения нейронной сети был использован набор изображений лиц в маске и без масок. Общее количество изображений в датасете составляет 1376 штук.

Для разработки системы распознавания, был выбран метод обнаружения лиц Виолы-Джонса. Этот метод был разработан в 2001 году, но благодаря своей высокой скорости, а также крайне низкой вероятности ложного обнаружения лица до сих пор является одним из основных методов поиска объектов на изображении.

Основные принципы, на которых основана работа данного метода:

- Представление изображения в интегральном виде;
- Поиск лиц с помощью признаков Хаара;
- Каскадная классификация;

В расширенном методе виолы-Джонса, который используется в данном проекте, используются дополнительные признаки Хаара. Каскады Хаара представляют собой прямоугольные области, которые составлены из нескольких соседних прямоугольных областей, отмеченных как светлая или темная, примеры таких дополнительных признаков представлены на рисунке 16.

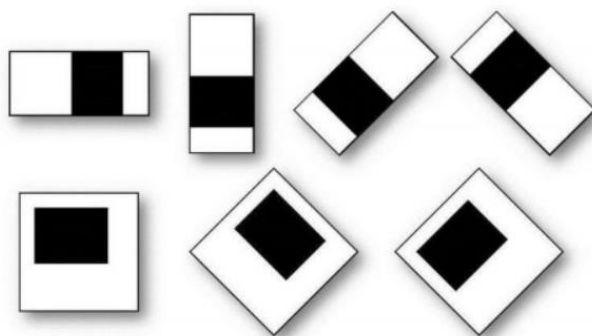


Рисунок 16 – Дополнительные признаки Хаара

Для того чтобы обучить модель нейросетевой детектор маски на лице, необходимо разбить ядро программы на две отдельные фазы, каждая со своим соответствующим этапом приведенные на рисунке 17:

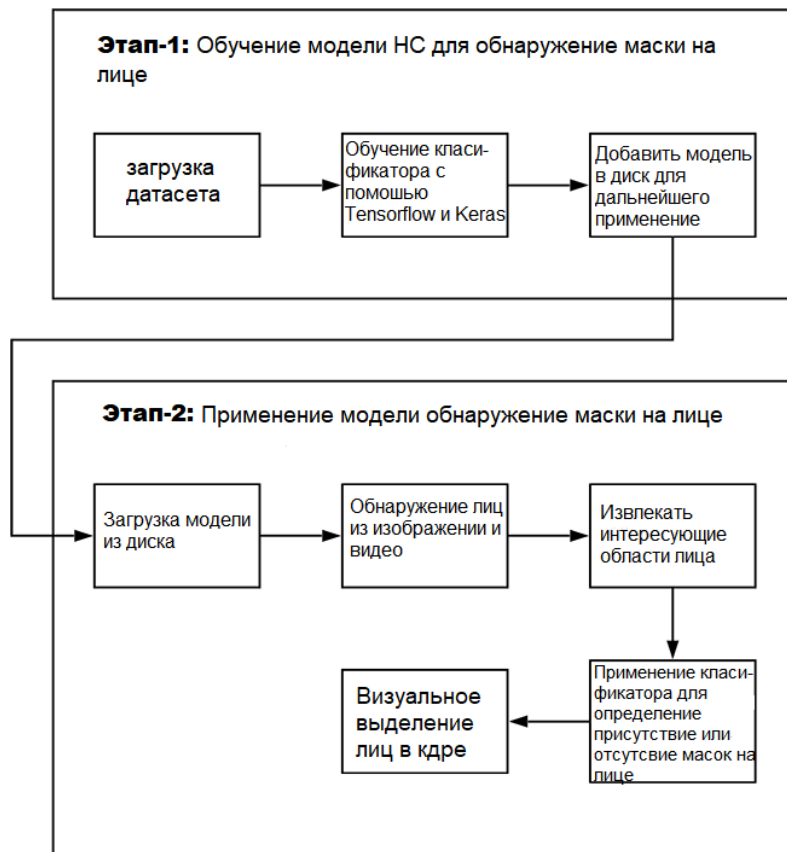


Рисунок 17. Алгоритм обработки кадров видеопотока для обнаружения лиц и маски на лице.

Обучение модели сверточной нейронной сети: на данном этапе, акцентируем внимание на загрузке нашего набора данных из памяти компьютера для обучения модели обнаружения маски на лице (с использованием Keras / TensorFlow), а затем сохраняем модель с лучшими показателями.

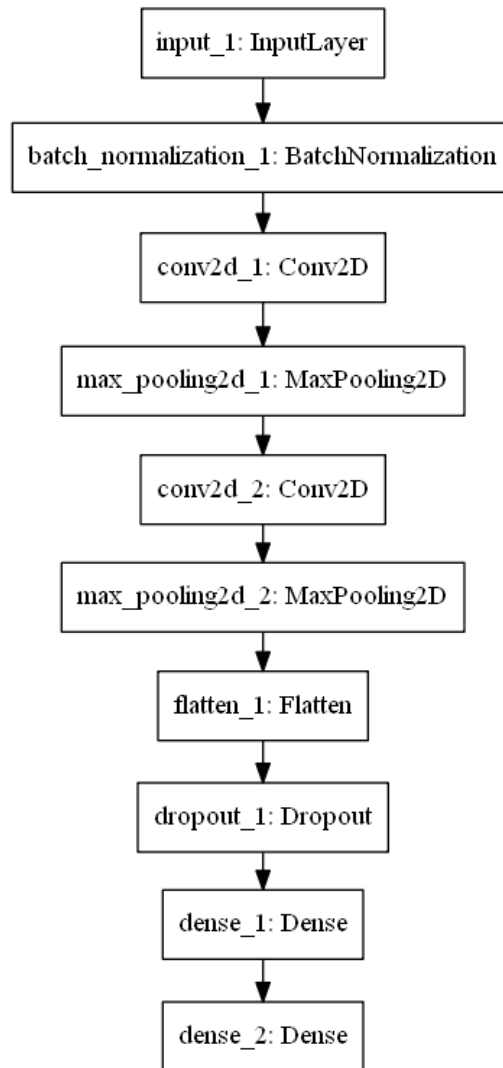


Рисунок 18. Модель сверточной нейронной сети

Развертывание: после того, как нейросетевой модель для детектирования маски обучен, переходим к загрузке детектора маски из диска для применения этапа обнаружения лиц, а затем классификации лица в защитной маске или же без защитной маски.

- Шаг 1. Загрузить и обработать данные для тренировки нашей нейронной сети.
- Шаг 2. Реализовать архитектуру свёрточной нейронной сети (вход: $224 * 224 * 3$).
- Шаг 3. Обучить модель на тренировочной выборке с помощью библиотеки, Tensorflow, Keras, Numpy и Matplotlib и сохранить модель для дальнейшего применение.

Шаг 4. Развертывать обученной модели для обработки кадров и для оценок признаков в видеопотоке в реальном времени.

Шаг 5. Протестировать модель на видеопотоке в реальном времени.

3. Результаты тестирования модели

Для того чтобы встроить видимые черты лица в 128-размерности, нам нужно обрезать обнаруженные лица, и для этого мы выбрали InceptionV3 за легкую архитектуру и точность. Он реализован с помощью Keras - нейросетевой библиотеки с открытым исходным кодом, работающей на базе библиотеки TensorFlow, разработанной для реализации задач компьютерного зрения. Такой подход является эффективным, так как нейросеть пытается обнаруживать признаки присутствия или отсутствия маски, только в тот фрагмент изображения, где действительно есть лица. Это позволяет улучшить скорость распознавание маски на лице.

Данный метод облегчает задачу распознавания, поскольку для сравнения признаков, используемые предварительно обученную модель глубокого обучения, требуется много времени, и нам нужно выполнить эту задачу за миллисекунды. По словам Флориана Шроффа, Дмитрия Калениченко и Джеймса Филбина (2015, 815-816), FaceNet использует ГСНС DCNN (глубокую сверточную нейронную сеть), которая обучена вычислять диапазон между вложениями, соответствующими сходству лиц [14].

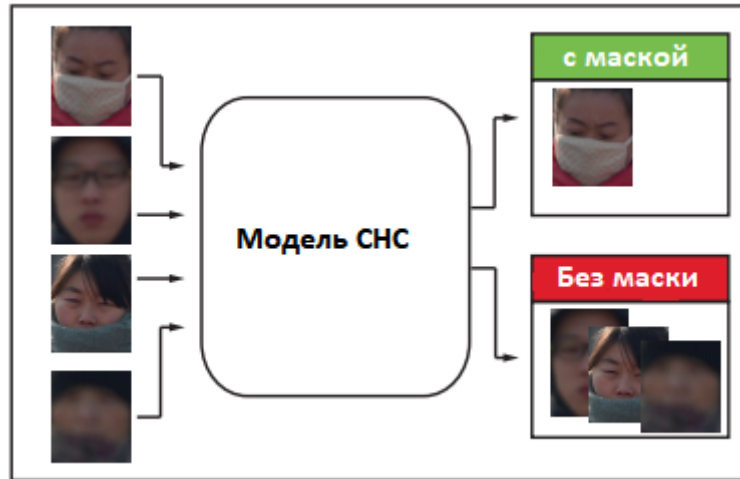


Рисунок 19 – Изображения лиц для обнаружения масок

В датасете использовали тройную функцию потерь для вычисления сходства между двумя встраиваемыми гранями. Поскольку у нас есть три входа для сети, а идея использования этой функции потерь состоит в том, что объект привязки должен быть относительно похож на положительный объект по сравнению с отрицательным объектом, и далее приведен метод для вычисления этого сравнение:

$$\mathcal{L}(x, y, z) = \max (\|f(x) - f(y)\|^2 - \|f(x) - f(z)\|^2 + \theta)$$

Где x обозначает объект привязки, y - положительный объект, а z - отрицательный объект. $F()$ обозначают функцию, которая встраивает изображение в 128-мерное изображение. И θ представляют собой границу между положительной и отрицательной частью, что означает различительное значение между парами изображений.

На этапе кодирования мы нормализовали встроенное лицо, что означает масштабирование значений, измеренных в другом диапазоне, до стандартной шкалы и нормализацию наших векторов, мы использовали библиотеку `scikits-learn`, которая имеет различные алгоритмы классификации, регрессии и кластеризации.

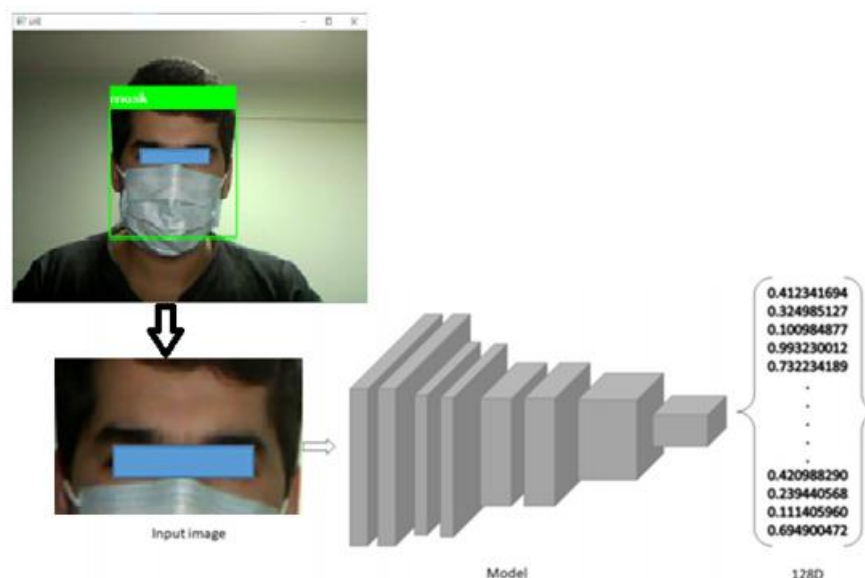


Рисунок 20 – Преобразование изображение лица в 128-мерный вектор

После выполнения предварительной работы по получению и подготовки данных для обучения модели, была реализована архитектура и обучена модель сверточная нейронная сеть.

Результаты обучения приведен на рисунке 21. Можно заметить, что модель хорошо обучилась и достигла точности распознавания, равной 98%.

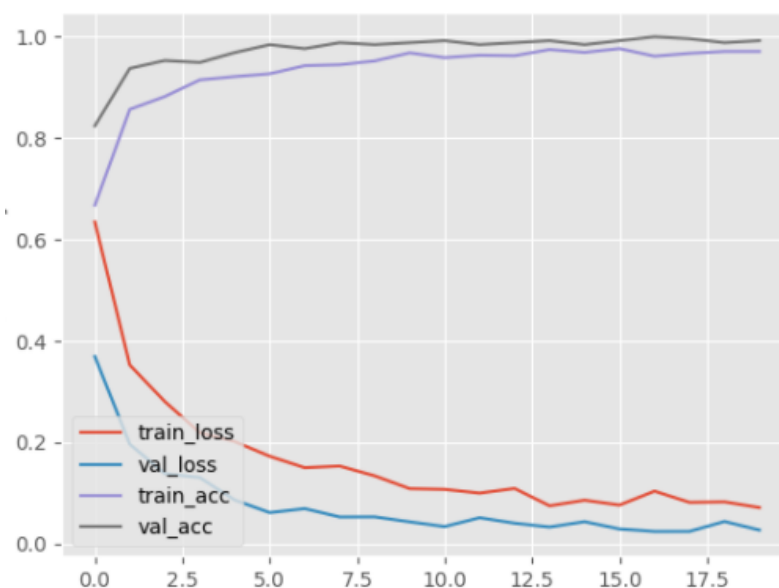


Рисунок 21. Результат проведение обучение модели.

При запуске программы и получении доступа к камере, открывается отдельное окно, в котором выводятся результаты распознавания лиц. После этапа распознавания лиц производится детектирование защитной медицинской маски на лице, а результат детектирования выводится на экране (рисунок 22).



Рисунок 22. Результат тестирования программы для лиц без защитной маски.

Убедившись в работоспособности программы для распознавания лиц без защитной маски, было произведено тестирование на лице человека в защитной маске (рисунок 23).

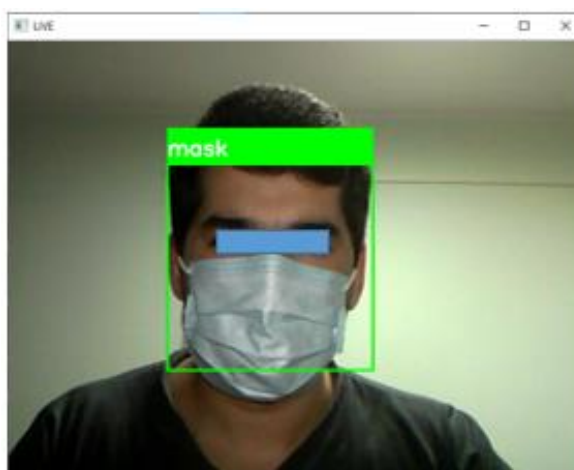


Рисунок 23. Результат работы программы на лицах в защитной маске

В результате тестирования были получены ожидаемые результаты, заключающиеся в том, что применение обученных сверточных нейронных сетей, позволяет обнаруживать маски, на распознанных лицах. Достигнутая точность распознавания масок составляет 98%, позволяя достаточно хорошо определить носит ли человек защитную маску или нет.

3.1. Поворот вокруг вертикальной оси

Использовались видео, в котором расположение лиц был изменен от под разными углами. Поворот головы вокруг вертикальной оси в диапазоне $[-60; +60]$ градусов с шагом в 15 градусов. Результаты работы первого этапа:

Угол, градусы	Точность распознанных лиц в маске (1/100%)
-60	0.82
-45	0.86
-30	0.92
-15	0.96
0	0.98
+15	0.95
+30	0.93
+45	0.88
+60	0.85

Можно видеть, что при 60 градусах, алгоритм распознается с хороших показателей равной 86% точности, при 45 - больше 88%. Таким образом, алгоритм стабильно работает при угле, меньшем 60^0 . При использовании алгоритма, распозналось дополнительно одно изображение, содержащее лицо, при угле 75% и два - при угле в 65%, при этом при меньшем угле поворота изменений нет. Это улучшение незначительно, но говорит о том, что начальное предположение о возможности аппроксимации верно. Стоит отметить, что в

реальных условиях в систему попадает лиц с поворотом по вертикальной оси менее 30-35 градусов.

3.2. Поворот вокруг оси Z

Тест состоит в повороте лиц на 360 градусов инкрементами по 15 градусов. Результат: при поворотах вплоть до +/- 30 градусов лицо обнаруживается.

3.3. Критерии принятия решения модели

Исследуем, чему научилась обученная модель, чтобы понять, что способствовало принятию решений модели о наличии или отсутствия маски. Для этого, используем технику отображения активации классов, взвешенных по градиенту (Grad-CAM) [11]. Grad-CAM использует градиент оценки классификации относительно сверточных признаков, определенных сетью, чтобы понять, какие части изображения являются наиболее важными для классификации. Несколько образцов изображений, которые были правильно классифицированы, как содержащие лица без масок, показаны на рисунке. Как можно видеть, открытые черты лица, такие как нос и рот, больше всего способствуют принятию решения.

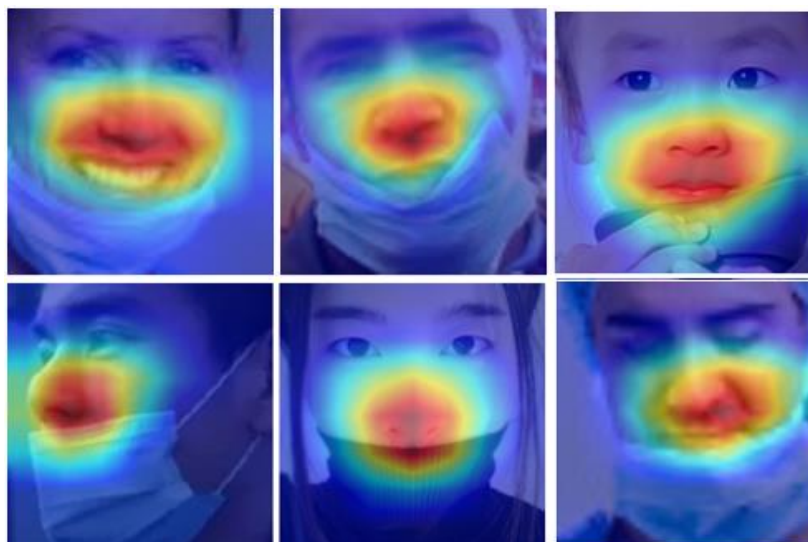


Рисунок 24. Правильно классифицированные лица, которые неправильно носят маску.

Исходя из полученных результатов, можно сделать вывод, что критерии для принятия решения модели соответствуют требованиям всемирной организации здравоохранения (ВОЗ) [12].

4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

4.1.Предпроектный анализ

Данная научно-исследовательская работа посвящена разработке методов и алгоритмов, входящих в программный комплекс, осуществляющий распознавание человеческих лиц с последующей выявление признаки присутствие или отсутствие защитной медицинской маски на лице. Данная работа реализовалась с использованием методов машинного обучения, компьютерного зрения и построенного по архитектуре искусственных нейронных сетей. Применения такой технологии можно найти в системах контроля доступа в помещениях, контроль доступа с помощью лиц, подсчет посещаемость групп лиц, проверка безопасности на вокзалах и т.д. Более конкретная сфера применения данной технологии — это обнаружение групп лиц, которые не соблюдают карантинных мер, социальное дистанцирование и не носят защитные медицинские маски, следовательно, разрешить или запретить им доступ в частных или общественных местах.

С точки зрения актуальности исследования в рамках социальной направленности результаты настоящего исследования позволят в значительной степени упростить и ускорить процесс контроля доступа или допуска в определенной территории или общественных местах таких как: метро, аэропорт, университет и т.д.

Для эффективного использования научного потенциала научно-исследовательского проекта необходимо не только провести исследовательскую работу, но и провести анализ исследования с точки зрения финансового менеджмента. Задачами данного раздела являются:

1. Определение перспективности технологии с помощью технологии QuaD
2. Определение сильных и слабых сторон, а также возможностей и угроз с помощью SWOT-анализа
3. Оценка готовности проекта к коммерциализации
4. Определение целей и результата научно-исследовательского проекта

5. Планирование работы, распределение задач и определение трудоемкости
6. Формирование бюджета научно-исследовательского проекта
7. Проведение анализа рисков

4.2.Технология QuaD

В рамках данной работы анализ конкурентных технических решений невозможен ввиду отсутствия открытых данных о наличии и свойствах подобных разработок. Это связано с тем, что такие программные решения стали актуальными после возникновения пандемии COVID-19 и находятся в стадии разработки, для нужд конкретной организаций либо являются частью более сложной системы идентификация.

Технология QuaD представляет собой гибкий инструмент измерения характеристик, описывающих качество новой разработки и её перспективность на рынке и позволяющие принимать решение целесообразности вложения денежных средств в научно-исследовательский проект.

- a) Энергоэффективность. Характеристика проявляется в потреблении энергии устройствами. Помимо алгоритма для самой системы необходима ЭВМ, для обработки входящих сигналов. У разрабатываемого метода для обнаружения лиц и распознавания маски, необходим только ноутбук или одноплатный компьютер.
- b) Помехоустойчивость. Данная характеристика определяется способностью справляться с шумами, создаваемыми окружающими источниками света
- c) Надежность. Модель сверточной нейросети был обучен на большом количестве качественных данных, это увеличить надежность работы алгоритма.

- d) Унифицированность. Система состоит из разных компонентов работающих для каждого отдельного этапа обнаружения и распознавания, которые были приведены к единообразной системе.
- e) Производительность алгоритма. Данная характеристика определяется способностью быстрого действия алгоритма при обнаружении и распознавании.
- f) Пользовательский интерфейс. Система имеет простой и интуитивно понятный интерфейс.
- g) Безопасность. Так как система в основном работает локально, это позволяет избегать опасности взлома системы.
- h) Функциональная мощность. Параметр определяется количеством распознаваемых лиц за раз. Разрабатываемый алгоритм за счет архитектуры нейронной сети, способен распознавать несколько лиц в кадре.
- i) Простота эксплуатации. Система способна работать на разных архитектурах ЭВМ, даже на одноплатных компьютерах размером с банковскую карту. Данная характеристика означает, что систему можно развертывать на разных операционных системах.
- j) Прозрачность кода. Данная характеристика определяется качеством и уникальностью кода, который был написан для разрабатываемой системы.

Составленная таблица QuaD представлена в таблице 1.

Таблица 1 – Оценочная карта технологии QuaD

Критерий оценки	Вес критерия	Баллы	Максимальный балл	Относительное значение	Средневзвешенное значение
Энергоэффективность	0,15	90	100	0,9	0,135
Помехоустойчивость	0,1	90	100	0,9	0,09
Надежность	0,15	95	100	0,95	0,1425
Унифицированность	0,05	80	100	0,8	0,04

Производительность алгоритма	0,05	65	100	0,65	0,0325
Пользовательский интерфейс	0,05	25	100	0,25	0,0125
Безопасность	0,1	100	100	1	0,1
Потребность в ресурсах памяти	0,1	75	100	0,75	0,075
Функциональная мощность	0,05	80	100	0,8	0,04
Простота эксплуатации	0,1	95	100	0,95	0,095
Качество интеллектуального интерфейса	0,05	80	100	0,8	0,04
Прозрачность кода	0,05	100	100	1	0,05
Итого	1				0,8525

Значение качества по технологии QuaD составило 0,8525. Можно сказать, что подобная разработка может считаться перспективной, так как у рассматриваемой разработки высокие показатели по всем наиболее важным критериям, таким как надежность, функциональная мощность и т.д.

4.3.SWOT-анализ

SWOT – Strengths (сильные стороны), Weaknesses (слабые стороны), Opportunities (возможности) и Threats (угрозы) – это комплексный анализ научно-исследовательского проекта. Такой анализ применяют для исследования внешней и внутренней среды проекта. Первый этап заключается в описании сильных и слабых сторон проекта, в выявлении возможностей и угроз для реализации проекта, которые проявились или могут появиться в его внешней среде. Составленная матрица SWOT представлена в таблице 2.

Таблица 2 – SWOT-анализ

<p>Сильные стороны научноисследовательского проекта:</p> <p>С1. Наличие опыта работы с методами глубокого обучения.</p> <p>С2. Достаточный уровень развития технологий компьютерного зрения для решения задачи.</p> <p>С3. Большое количество библиотек с открытым исходным кодом.</p> <p>С4. Наличие промышленного опыта разработки ПО.</p>	<p>Слабые стороны научноисследовательского проекта:</p> <p>Сл1. Высокая трудоемкость методов исследования.</p> <p>Сл2. Высокие требования к вычислительным ресурсам для обучения нейронной сети для алгоритма.</p> <p>Сл3. Дефицит размеченных данных для обучения нейронной сети.</p> <p>Сл4. Нехватка сотрудников для исследования.</p>
<p>Возможности:</p> <p>В1. Возможность применения наработок по исследованию в других областях.</p> <p>В2. Возможность использовать бесплатные вычислительные мощности, предоставляемые для студентов и исследователей.</p>	<p>Угрозы:</p> <p>У1. Невозможность внедрения из-за отрицательных результатов.</p> <p>У2. Несвоевременное финансирование исследования.</p> <p>У3. Увеличение сроков исследования из-за нехватки вычислительных ресурсов.</p>

В рамках SWOT анализа были рассмотрены по четыре слабых и сильных сторон проекта, а также возможности и угрозы. Указаны стратегии, которые будут ликвидировать слабые стороны разработки и бороться с угрозами.

Далее выявляются соответствия сильных и слабых сторон проекта условиям внешней среды. Это необходимо для оценки необходимости какихлибо изменений проекта. В таблице 3 представлены соответствия сильных сторон проекта и возможностей его реализации.

Таблица 3 – Интерактивная матрица проекта

Сильные стороны исследования					
Возможности исследования		Си1	Си2	Си3	С4
	В1	+	+	+	–
	В2	–	–	–	+

Из таблицы 3 следует, что сильные стороны исследования соответствуют возможностям по следующим параметрам:

1. В1, С1, С3, С3
2. В2, С1

Возможность В1 соответствует сильным сторонами С1С3С4, т.к. возможность применения наработок в других областях зависит от наличия опыта работы с методами глубокого обучения, текущего уровня развития технологий, а также от наличия соответствующих библиотек с открытым исходным кодом.

Возможность В2 соответствует сильной стороной С4, т.к. наличие опыта разработки ПО позволяет с легкостью использовать бесплатные вычислительные ресурсы для академических целей.

В таблице 4 представлены соответствия слабых сторон проекта и возможностей его реализации.

Таблица 4 – Слабые стороны исследования и возможности

Слабые стороны исследования					
Возможности исследования		Си1	Си2	Си3	С4
	В1	+	+	+	–
	В2	+	+	–	–

Из таблицы 4 следует, что сильные стороны исследования соответствуют возможностям по следующим параметрам:

1. В1, Сл1, Сл2, Сл3
2. В2, Сл2, Сл4

Возможность В1 взаимодействует со слабыми сторонами Сл1Сл2Сл3, т.к. трудоемкость исследования, высокие требования к вычислительным ресурсам и нехватка размеченных данных могут помешать использовать наработки в других областях.

Возможность В2 взаимодействует со слабыми сторонами Сл1Сл2, т.к. трудоемкость исследования и высокие требования к вычислительным ресурсам не позволят полноценно использовать бесплатные вычислительные ресурсы.

В таблице 5 представлены соответствия слабых сторон проекта и возможностей его реализации.

Таблица 5 – Сильные стороны исследования и угрозы

Сильные стороны исследования					
Угрозы исследования		Си1	Си2	Си3	С4
	У1	–	–	–	–
	У2	–	–	+	+
	У3	–	–	–	–

Из таблицы 5 следует, что сильные стороны исследования соответствуют угрозам по следующим параметрам:

1. У2 С2 С4

Угроза У2 взаимодействует с сильными сторонами С3С4, т.к. наличие опыта разработки ПО позволит не привлекать дополнительных специалистов, а наличие библиотек с открытым исходным кодом позволит сэкономить на покупке коммерческого ПО, что даст возможность начать исследование не дожидаясь финансирования.

В таблице 6 представлены соответствия слабых сторон проекта и возможностей его реализации.

Таблица 6 – Слабые стороны исследования и угрозы

Слабые стороны исследования					
Угрозы исследования		Си1	Си2	Си3	С4
	У1	–	–	–	–
	У2	–	+	–	–
	У3	–	+	–	–

Из таблицы 6 следует, что сильные стороны исследования соответствуют угрозам по следующим параметрам:

1. У2Сл2
2. У3Сл2

Угрозы У2У3 взаимодействует со слабой стороной С2, т.к. высокие требования к вычислительным ресурсам могут сильно растянуть проект во времени с учетом угроз несвоевременного финансирования.

Анализ показал, что главными угрозами проекту являются запоздание исследования в связи с нехваткой вычислительных ресурсов и несвоевременное финансирование исследования. Слабыми сторонами проекта высокая трудоемкость методов исследования, высокие требования к вычислительным ресурсам для обучения нейронной сети для алгоритма, дефицит размеченных данных для обучения нейронной сети, нехватка сотрудников для исследования. Тем не менее, проект обладает и сильными сторонами, позволяющими частично компенсировать слабые стороны, а также уменьшить негативное влияние возможных угроз:

- а) Использование сервисов облачных вычислений больших компании таких как Google и Amazon, компенсирует нехваткой вычислительных ресурсов для обучения моделей неросети.

- b) Повысить точность и качество работы системы, можно достигать путём оптимизации работы алгоритма и улучшение качество обучаемых данных.
- c) Данную систему можно объединить как часть продукта большой компании. Это позволяет привлекать финансирование, тем самым расширить сфера применения системы и масштабировать количество сотрудников для исследования.

4.4. Оценка готовности проекта к коммерциализации

Для научно-исследовательского проекта важно оценить степень его готовности к коммерциализации, а также оценить уровень собственных знаний для её проведения. По результатам такого анализа можно сделать вывод о готовности научно-исследовательского проекта к коммерциализации, а также о необходимости привлечения иных специалистов в команду проекта (таблица 7).

Таблица 7 – Оценки степени готовности научноисследовательского проекта к коммерциализации.

№ п/п	Наименование	Ком- мерция	Компе- тенции
1	Определен имеющийся научно-технический задел	4	4
2	Определены перспективные направления коммерциализации научно-технического задела	4	4
3	Определены отрасли и технологии (товары, услуги) для предложения на рынке	4	4
4	Определена форма для представления на рынок	4	4
5	Определены авторы и осуществлена охрана их прав	2	3
6	Оценена стоимость интеллектуальной собственности	2	2
7	Проведены маркетинговые исследования рынков сбыта	3	3
8	Разработан бизнес-план коммерциализации разработки	2	3
9	Определены пути продвижения разработки на рынок	3	4
10	Разработана стратегия (форма) реализации разработки	4	4
11	Проработаны вопросы международного сотрудничества и выхода на зарубежный рынок	1	1

12	Проработаны вопросы использования услуг инфраструктуры поддержки, получения льгот	2	2
13	Проработаны вопросы финансирования коммерциализации научной разработки	3	4
14	Имеется команда для коммерциализации разработки	2	3
15	Проработан механизм реализации научно-исследовательского проекта	4	4
	ИТОГО БАЛЛОВ	44	49

Итоговые значения проработанности научного проекта и знания у разработчика 44, что говорит о том, что некоторые аспекты проекта практически не были проработаны, а именно охрана прав, а также анализ рынков сбыта.

Можно сделать вывод, что для дальнейшей коммерциализации научно-исследовательского проекта необходимо вовлечение сторонних специалистов в этой области и более глубокая проработка плана коммерциализации. Успех продвижения товара на рынок во многом зависит от правильности выбора метода коммерциализации. Для данного научно-исследовательского проекта наиболее приемлемым методом коммерциализации является создание собственного предприятия для реализации конечного продукта и его продажи частным и государственным организациям.

Группа процессов инициации состоит из процессов, направленных на определение нового научно-исследовательского проекта. В рамках процессов инициации определяются цели, а также заинтересованные стороны проекта.

В данном разделе приводится информация о заинтересованных сторонах научно-исследовательского проекта, иерархии целей проекта и критериях достижения целей.

Таблица 6 содержит информацию о заинтересованных сторонах научно-исследовательского проекта.

Таблица 6 – Заинтересованные стороны научно-исследовательского проекта

Заинтересованные стороны	Ожидания
Отделение информационных технологий ТПУ	Научные публикации Защита магистерской диссертации
Государственные организации	Обеспечение безопасности людей
Компания пользователя	Облегчен процесс контроля за проектом компании
Научный руководитель, студент	Выполненная выпускная квалификационная работа
Научное сообщество	Алгоритм обнаружение лиц в защитные маски

В таблице 7 представлена информация об иерархии целей научно-исследовательского проекта и критериев их достижения.

Таблица 7 – Иерархия целей проекта и критерии их достижения

Цели проекта:	Разработать способ обнаружение лиц и распознавание защитных медицинских масок в лице
Ожидаемые результаты проекта:	Алгоритм для обнаружение лиц и распознавание защитных медицинских масок в лице на видеопотоке
Требования к результату проекта	Разработан способ обучения нейронной сети для точной распознавание масок
	Разработан алгоритм обнаружение лиц и распознавание защитных медицинских масок в лице с небольшими требованиями к вычислительным ресурсам для работы на персональных компьютерах
	Бесперебойная работа программных модулей проекта

	Формализованное описание работы программных модулей проекта
--	---

4.5. Планирование управления научно-техническим проектом

4.5.1. План проекта

Диаграмма Ганта – это тип столбчатых диаграмм (гистограмм), который используется для иллюстрации календарного плана проекта, на котором работы по теме представляются протяженными во времени отрезками, характеризующимися датами начала и окончания выполнения данных работ.

График строится в виде таблицы 8 с разбивкой по декадам (10 дней) за период времени выполнения научного проекта. При этом работы на графике следует выделить различной штриховкой в зависимости от исполнителей, ответственных за ту или иную работу.

Таблица 8. Календарный план-график проведения НИОКР по теме

Ко д	Вид	Исп	Тк, ч. / Т раб	Продолжительность														
				Янв.			Февр.			Март			Апр.			Май		
				1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
1	Выбор направления исследования	Р, С	20 / 2.5															
2	Описание требований	С	15 / 2															
3	Патентный поиск	С	5 / 1															

4	Составление технического задания	С	15 / 2														
5	Изучение литературы	С	20 / 2.5														
6	Изучить систему	С	50 / 6														
7	Анализ предметной области	С	12 / 1														
8	Подготовка данных	С	15 / 1														
9	Выбор метода обнаружение лиц	С	25 / 3														
10	Выбор архитектуры нейронной сети	С	25 / 3														
11	Согласование выполненной работы с научным руководителе м	Р	4 / 1														
		С	25 / 3														
12	Написание документации	С	25 / 3														
13	Оптимизирова ть работы алгоритма	С	51 / 6														
14	Выполнение других частей работы (финансовый менеджмент, социальная ответственнос ть)	С	21 / 2.5														
15	Проверка работы	Р	20 / 2.5														

■ – Руководитель(Р)

■ – Студент (С)

Итого руководитель потратит на проект около 5 дней, а студент примерно 39.

4.5.2. Бюджет научного исследования

При планировании бюджета научного исследования должно быть обеспечено полное и достоверное отражение всех видов планируемых расходов, необходимых для его выполнения. В процессе формирования бюджета, планируемые затраты группируются по статьям, представленным в таблице.

4.5.2.1. Специальное оборудование для научных работ

В данную статью включают все затраты, связанные с приобретением специального оборудования (приборов, контрольно-измерительной аппаратуры, стендов, устройств и механизмов), необходимого для проведения работ по конкретной теме. Определение стоимости спецоборудования производится по действующим прейскурантам, а в ряде случаев по договорной цене.

Для разработки нынешней системы требуется обычное оборудование в виде персонального компьютера и купленных серверов на время жизненного цикла программного продукта. Среда и средство разработки, программный софт и другие комплектующие, нужные для разработки, распространяются бесплатно и не требуют дополнительных затрат (таблица 9).

Таблица 9 – Расчет затрат на «Спецоборудование для научных работ»

№	Наименование оборудования	Кол-во	Цена, руб.	Стоимость, руб.
1.	Персональный компьютер	1	-	-
2.	Среда разработки PyCharm	1	5 800	5 800
3.	Google Colab Pro	1	7 890	7 890

4.	ИК-камера	1	6 950	6 950
Итого				20 640

4.5.2.2. Основная заработная плата

В настоящую статью включается основная заработная плата научных и инженерно-технических работников, рабочих макетных мастерских и опытных производств, непосредственно участвующих в выполнении работ по данной теме.

Должность руководителя – доцент, к.т.н. – 35 111.5 рублей в месяц

Должность инженера – студент – 22 695,68 рублей в месяц

Расчет основной заработной платы сводится в таблице 10.

Таблица 10 – Расчет основной заработной платы

№ п/п	Наименование этапов	Исп	Трудоемкость, чел.-дн.	Зарплата, чел.-дн., руб	Всего руб.
1		Р	4	35 111.5	35 111.5
2		С	39	22 695.68	22 695.68

$$C_{зп} = Z_{осн} + Z_{доп},$$

где $Z_{осн}$ – основная заработная плата;

$Z_{доп}$ – дополнительная заработная плата.

Основная заработная плата $Z_{осн}$ руководителя рассчитывается по следующей формуле:

$$Z_m = Z_{дн} \cdot T_{раб}$$

где $T_{раб}$ – продолжительность работ, выполняемых научно-техническим работником, рабочие дни. (таблица 14);

$Z_{дн}$ – среднедневная заработная плата работника, руб.

Значит, для руководителя:

$$Z_m = 35\,111.5 \cdot 1,3 = 45\,645 \text{ рублей}$$

Для инженера:

$$З_{\text{м}} = 22\,695.68 * 1,3 = 29\,504 \text{ рублей}$$

Среднедневная заработная плата рассчитывается по формуле:

$$З_{\text{дн}} = (З_{\text{м}} \cdot М) / F_{\text{д}}$$

где $З_{\text{м}}$ – месячный должностной оклад работника, руб

$М$ – количество месяцев работы без отпуска в течение года:

при отпуске в 45 раб. дней $М=10,4$ месяца, 6-дневная неделя;

$F_{\text{д}}$ – действительный годовой фонд рабочего времени научно-технического персонала (в рабочих днях) (таблица 13). Тогда,

Для руководителя:

$$З_{\text{дн}} = \frac{45\,645 * 10,4}{254} = 1\,869 \text{ рублей}$$

Для инженера:

$$З_{\text{дн}} = \frac{22\,695.68 * 11,2}{217} = 1\,171 \text{ рублей}$$

Баланс рабочего времени представлен в таблице 11.

Таблица 11 – Баланс рабочего времени

Показатели рабочего времени	Руководитель	Инженер
Календарное число дней	365	365
Количество нерабочих дней		
- выходные дни	52	82
- праздничные дни	11	14
Потери рабочего времени		
- отпуск	56	24
- невыходы по болезни	—	—
Действительный годовой фонд рабочего времени	254	217

Таблица 12 – Результаты расчета основной заработной платы

Исполнители	$З_{\text{б}}$, руб.	$k_{\text{р}}$	$З_{\text{м}}$, руб	$З_{\text{дн}}$, руб.	$T_{\text{р}}$, дн.	$З_{\text{осн}}$, руб.
Руководитель	35 111.5	1.3	45 645	1 869	4	7 476

Инженер	22 695.68	1.3	29 504	1 171	39	45 669
Итого по статье $Z_{\text{осн}}$:			53 145			

4.5.2.3. Дополнительная заработная плата научно-производственного персонала

Дополнительная заработная плата включает оплату за непроработанное время (очередной и учебный отпуск, выполнение государственных обязанностей, выплата вознаграждений за выслугу лет и т.п.) и рассчитывается исходя из 10-15% от основной заработной платы, работников, непосредственно участвующих в выполнение темы:

$$Z_{\text{доп}} = k_{\text{доп}} * Z_{\text{осн}} \quad (4.9)$$

где $Z_{\text{доп}}$ – дополнительная заработная плата, руб.;

$k_{\text{доп}}$ – коэффициент дополнительной зарплаты ($k_{\text{доп}} = 0,1$);

$Z_{\text{осн}}$ – основная заработная плата, руб.

Для руководителя:

$$Z_{\text{доп}} = 7\,476 * 0,1 = 747.6 \text{ рублей}$$

Для инженера:

$$Z_{\text{доп}} = 45\,669 * 0,1 = 4\,566.9 \text{ рублей}$$

В таблице 13 приведен расчёт основной и дополнительной заработной платы.

Таблица 13 – Заработная плата исполнителей ВКР, руб

Заработная плата	Руководитель	Инженер
Основная зарплата	7 476	45 669
Дополнительная зарплата	747.6	4 566.9
Зарплата исполнителя	8 223.6	50 235.9
Итого	58 459	

4.5.2.4. Отчисления на социальные нужды

Статья включает в себя отчисления во внебюджетные фонды.

$$C_{\text{внеб}} = k_{\text{внеб}} \cdot (Z_{\text{осн}} + Z_{\text{доп}}) = 0,3 \cdot (7\,476 + 747,6) = 2\,467,1 \text{ руб.}$$

Для инженера:

$$C_{\text{внеб}} = k_{\text{внеб}} \cdot (Z_{\text{осн}} + Z_{\text{доп}}) = 0,3 \cdot (45\,669 + 4\,566,9) = 15\,070,7 \text{ руб.}$$

где $k_{\text{внеб}}$ – коэффициент отчислений на уплату во внебюджетные фонды (пенсионный фонд, фонд обязательного медицинского страхования и пр.).

Для руководителя – 2 467,1 руб.

Для инженера – 15 070,7 руб.

Итого 17 538 рублей

4.5.2.5. Накладные расходы

В эту статью относятся расходы по содержанию, эксплуатации и ремонту оборудования, производственного инструмента и инвентаря, зданий, сооружений и др. В расчетах эти расходы принимаются в размере 16 % от суммы основной заработной платы научно-производственного персонала данной научно-технической организации.

Накладные расходы составляют 16% от суммы основной и дополнительной заработной платы, работников, непосредственно участвующих в выполнение темы.

Расчет накладных расходов ведется по следующей формуле:

$$C_{\text{накл}} = k_{\text{накл}} \cdot (Z_{\text{осн}} + Z_{\text{доп}})$$

где $k_{\text{накл}}$ – коэффициент накладных расходов.

$$C_{\text{накл}} = 0,16 \cdot 58\,459 = 9\,353 \text{ руб.}$$

4.5.2.6. Формирование бюджета затрат научно-исследовательского проекта

Рассчитанная величина затрат научно-исследовательской работы является основой для формирования бюджета затрат проекта, который при формировании договора с заказчиком защищается научной организацией в качестве нижнего предела затрат на разработку научно-технической продукции.

Таблица 14 – Бюджет затрат НТИ

Сырье	Оборудован	Осн зп	Доп зп	Накл рас	Отчисл на соц	Итого
-	30 258	53 145	5 314	9 353	17 538	115 608

В результате было получено, что бюджет на разработку НТИ составит 115 608 руб.

4.6. Риски научно-исследовательского проекта

При разработке научно-исследовательского проекта следует понимать и учитывать возможные риски. Таблица 15 содержит результаты анализа возможных рисков.

Таблица 24 – Реестр рисков

№	Риск	Потенциальное воздействие	Вероятность наступления (1-5)	Влияние риска (1-5)	Уровень риска	Способы смягчения риска	Условия наступления
1	Кадровый риск	Отсутствие заинтересованных исполнителей проекта	3	4	Высокий	Повышение мотивации исполнителей проекта	Потеря интереса исполнителей к деятельности проекта
2	Технический риск	Потеря файлов проекта	1	5	Высокий	Регулярное создание резервных копий файлов проекта	Отказ используемого оборудования

3	Доступ к данным	Отсутствие данных для работы системы	1	4	Средний	Заключение официального договора на доступ к данным	Отсутствие размеченных данных
4	Недостаточная производительность алгоритма	Неконкурентоспособность разработки	1	3	Средний	Проведение сравнительного тестирования, прототипирование	Ошибки при постановке задачи, недостаточный анализ качества разработки и ее перспективности на рынке

Из анализа реестра рисков можно заметить, что первый и второй вид рисков встречается во многих проектах. Риск же отсутствия доступа к размеченным данным во время выполнения данного научно-исследовательского проекта существенен для реализации, однако маловероятен. Также существует риск, что разработанный алгоритм будет недостаточно эффективен, но он также маловероятен при должном тестировании в процессе разработки прототипа.

4.7. Выводы по разделу «финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Научно-исследовательская работа заключалась в разработке алгоритма обнаружение лиц и распознавание лиц в защитные маски. Исследование проводилось при помощи ЭВМ. Данный исследовательский проект является перспективным согласно расчету с использованием технологии QuaD. Согласно SWOT-анализу проект имеет достаточно сильных сторон для использования возможностей и противодействия угрозам проекта. Для коммерциализации научно-исследовательского проекта необходимо вовлечение сторонних специалистов в этой области и создание собственного предприятия для разработки продукта.

Сформированный SWOT-анализ, в котором балы описаны сильные и слабые стороны проекта, в выявлении возможностей и угроз для реализации проекта, для выявления соответствия и несоответствия была составлена интерактивная матрица проекта.

Проведена оценка готовности проекта к коммерциализации, которая показала, что перспективность разработки средняя.

В рамках процессов инициации определены внутренние и внешние заинтересованные стороны проекта с их ожиданиями от проекта, цели и результат проекта.

План проекта представлен на диаграмме Ганта, из которого видно какой исполнитель (студент или руководитель) какой вид работ осуществлял и в течении какого количества дней.

В бюджет инженерно-технического проекта занесена стоимость оборудования, которая составила 20 640 рублей. Была рассчитана основная и дополнительная заработная плата исполнителей проекта, сумма которых составила 58 459 рублей. Вместе с отчислениями на социальные нужды и накладными расходами бюджет проекта получился равным 115 608 рублей.

Все рассмотренные риски научно-исследовательского проекта являются существенными для его реализации, однако вероятность их наступления достаточно мала.

5. Социальная ответственность

Введение

Магистерская диссертация посвящена разработке методов и алгоритмов, входящих в программный комплекс, осуществляющий распознавание человеческих лиц с последующим выявлением признаков присутствия или отсутствия защитной медицинской маски на лице. Данная работа реализовалась с использованием методов машинного обучения, компьютерного зрения на основе архитектуры искусственных нейронных сетей. Применения такой технологии можно найти в системах контроля доступа в помещениях, контроль доступа с помощью лиц, подсчет посещаемость групп лиц, проверка безопасности на вокзалах и т.д. Более конкретная сфера применения данной технологии — это обнаружение групп лиц, которые не соблюдают карантинных мер, социальное дистанцирование и не носят защитные медицинские маски, следовательно, разрешить или запретить им доступ в частных или общественных местах.

С точки зрения актуальности исследования в рамках социальной направленности результаты настоящего исследования позволят в значительной степени упростить и ускорить процесс контроля доступа или допуска в определенной территории или общественных местах таких как: метро, аэропорт, университет и т.д.

В данном разделе будут рассмотрены вредные и опасные факторы, которые оказывают влияние на рабочую деятельность персонала, рассмотрены воздействия разрабатываемой программы на окружающую среду, исследованы правовые и организационные вопросы обеспечения безопасности при чрезвычайных ситуациях.

5.1. Правовые и организационные вопросы обеспечения безопасности

В данном разделе рассмотрены основные нормативные документы, выполнение которых необходимо для безопасного исполнения рабочих обязанностей диспетчером, который проводит большую часть времени за компьютером.

Регулирование отношений между работником и работодателем, осуществляется трудовым кодексом РФ [21].

Продолжительность рабочего дня не должна превышать 40 часов в неделю. В течение рабочего дня работнику должен быть предоставлен перерыв для отдыха и питания продолжительностью не более двух часов и не менее 30 минут. Всем работникам предоставляются выходные дни.

ТОИ Р-45-084-01 Типовая инструкция по охране труда при работе на персональном компьютере регулируют порядок организации рабочего процесса для лиц, труд которых связан с компьютерной техникой [37]. Требования документа распространяются на персональные компьютеры, периферийные устройства (клавиатуры, принтеры, модемы, блоки бесперебойного питания и т.д.), а также на видео-дисплейные терминалы всех типов.

Организация рабочего места должна способствовать максимально удобной и безопасной работе сотрудника. Большое значение для профилактики статических физических перегрузок имеет правильная организация рабочего места человека, работающего с ПЭВМ. Рабочее место должно быть организовано в соответствии с требованиями стандартов, технических условий и (или) методических указаний по безопасности труда [22]. Конструкция рабочего места и взаимное расположение всех его элементов (сиденье, органы управления, средства отображения информации и т.д.) должны соответствовать антропометрическим, физиологическим и психологическим требованиям, а также характеру работы [23].

Основное внимание необходимо уделить делению на рабочие зоны так, как указано на рисунках 4.2 и 4.3, чтобы каждый сотрудник мог работать максимально эффективно, как отдельно так при необходимости работать в команде. Эргономика рабочего места напрямую влияет на производительность труда [22]. Согласно требованиям:

- расстояние между рабочими столами с видеомониторами должно быть не менее 2,0 м, а расстояние между боковыми поверхностями видеомониторов – не менее 1,2 м;
- экран видеомонитора должен находиться от глаз пользователя на расстоянии от 600 до 700 мм, но не ближе 500 мм с учетом размеров алфавитно-цифровых знаков и символов;

На рисунках 4.1 и 4.2 схематично представлены требования к рабочему месту и пространству вокруг работника.



Рисунок 4.1 – Организация рабочего места.

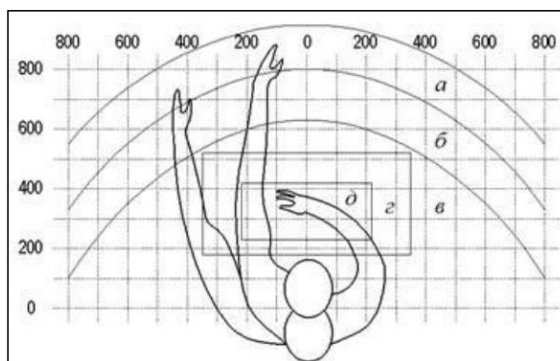


Рисунок 4.2 – Эргономические требования: а – зона максимальной досягаемости; б – зона досягаемости пальцев при вытянутой руке; в – зона

легкой досягаемости ладони; г – оптимальное пространство для грубой ручной работы; д – оптимальное пространство для точной ручной работы.

Оптимальное размещение предметов труда и документации в зонах досягаемости согласно [23]:

- дисплей размещается в зоне «а» (в центре);
- системный блок размещается в предусмотренной нише стола;
- клавиатура – в зоне «г/д»;
- «мышь» – в зоне «в» справа;
- документация, необходимая при работе – в зоне легкой досягаемости ладони – «б», а в выдвижных ящиках стола – редко используемая литература.

5.2. Производственная безопасность

Для обеспечения безопасности во время эксплуатации и разработки программы, необходимо провести анализ вредных и опасных воздействий на человека, которые могут возникать при разработке или эксплуатации проекта. Производственный фактор является вредным, в том случае если он приводит к заболеванию работника. В случае если его воздействие может привести к травме, то фактор является опасным.

В таблице 4.1 приведены возможные и опасные вредные факторы:

Таблица 4.1 – Возможные и опасные вредные факторы.

Источник фактора, наименования видов работ	Факторы (по ГОСТ 12.0.003-2015)		Нормативные документы
	Вредные	Опасные	
Работы за ПЭВМ	1. Отклонение показателей	1. Поражение электрическим током.	ГОСТ 22269-76 [3]; СанПиН 2.2.4.548-96 [5]; СанПиН 1.2.3685-21 [6];

	микроклимата от нормы. 2. Превышение уровня шума. 3. Отсутствие или недостаток освещения. 4. Психофизиологические факторы (монотонность труда, нервно-психические перегрузки, перенапряжение зрительных анализаторов).	2. Короткое замыкание. 3. Статическое электричество.	СП 52.13330.2016 [7]; ГОСТ 12.2.032-78 ССБТ [2]; СН 2.2.4/2.1.8.562-96 [9]; ГОСТ 12.1.030-81 Система стандартов безопасности труда (ССБТ). Электробезопасность. Защитное заземление. Зануление [10]; ГОСТ 26522-85 Короткие замыкания в электроустановках [11]; ГОСТ 12.1.038-82 Система стандартов безопасности труда (ССБТ). Электробезопасность. Предельно допустимые значения напряжений прикосновения и токов [12]; ГОСТ 12.1.004-91 [13].
--	---	---	---

5.2.1. Отклонение показателей микроклимата от нормы

Пониженная или же повышенная температура воздуха рабочей зоны считается вредным производственным фактором и является фактором микроклимата рабочей среды, параметры которого регулируются СанПиН 2.2.4.548-96 [25]. Он более характерен для рабочей среды программиста-разработчика.

Параметры, характеризующие микроклимат в производственных помещениях, являются:

- Температура воздуха (t , °C);
- Температура поверхностей (t , °C);
- Относительная влажность воздуха (ϕ , %);
- Скорость движения воздуха (v , м/с);

- Интенсивность теплового облучения (I , Вт/м²).

Во время работы с ПЭВМ в производственных помещениях происходит систематическое выделение тепла вычислительной техникой, а также вспомогательными устройствами и средствами освещения. По причине того, что оператор находится поблизости от источников выделения тепла, данный фактор является одним из вредных факторов производственной среды оператора ПЭВМ, а высокая температура воздуха способствует быстрому перегреву организма и повышению утомляемости.

Оптимальные нормы температуры, относительной влажности и скорости движения воздуха производственных помещений для работ, производимых сидя и не требующих систематического физического напряжения (категория Ia), приведены в таблице 7.2.

Для создания благоприятных условий труда и повышения производительности необходимо поддерживать оптимальные параметры микроклимата производственных помещений. Для этого должны быть предусмотрены следующие средства: центральное отопление, вентиляция (искусственная и естественная), искусственное кондиционирование.

Исходя из требований, рассматриваемых в данном разделе, в аудитории, в которой проводилось исследование поддерживается температура равная 20–21 °С, при относительной влажности в 55–58%. В помещении функционирует система принудительной вентиляции. Проводится ежедневная влажная уборка и систематическое проветривание помещения. В зимнее время в помещении предусмотрена система водяного отопления со встроенными нагревательными элементами и терморегуляторами.

Таблица 4.2 – Оптимальные параметры микроклимата в производственных помещениях пользователя ПЭВМ.

Период года	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность, %	Скорость движения воздуха, м/с
-------------	-------------------------	------------------------------	----------------------------	--------------------------------

Холодный	22-24	21-25	40-60	0,1
Теплый	23-25	22-26	40-60	0,1

5.2.2. Недостаточная освещенность рабочей зоны

Недостаточная освещенность рабочей зоны является вредным производственным фактором, возникающим при работе с ПЭВМ. Причиной недостаточной освещенности являются недостаточность естественного освещения, недостаточность искусственного освещения, пониженная контрастность.

Работа с компьютером подразумевает постоянный зрительный контакт с дисплеем ПЭВМ и занимает от 80% рабочего времени. Недостаточность освещения снижает производительность труда, увеличивает утомляемость и количество допускаемых ошибок, а также может привести к появлению профессиональных болезней зрения.

Существуют общие требования и рекомендации к организации освещения на рабочем месте:

- Искусственное освещение в помещениях для эксплуатации ПЭВМ должно осуществляться системой общего равномерного освещения.
- Рабочие места следует размещать таким образом, чтобы естественный свет падал преимущественно слева, а дисплеи монитора были ориентированы боковой стороной к световым проемам [2Ошибка! Источник ссылки не найден.].

Согласно СП 52.13330.2016 [27] в лаборатории, где происходит работа III зрительного разряда и подразряда г (работы высокой точности). В таблице 7.3 представлены нормативные показатели искусственного освещения при работах заданной точности.

Таблица 4.3 – Требования к освещению на рабочих местах для пользователей ПЭВМ.

Характеристика зрительной работы	Наименьший или эквивалентный размер объекта различения, мм	Разряд зрительной работы	Подразряд зрительной работы	Контраст объекта с фоном	Характеристика фона	Искусственное освещение		
						Освещенность, лк		
						При системе комбинированного освещения		При системе общего освещения
						всего	В том числе от общего	
Высокой точности	От 0,3 до 0,5	III	г	Средний, большой	Светлый, средний	400	200	200

Безопасность и здоровье условия труда в большой степени зависят от освещенности рабочих мест и помещений. Неудовлетворительное освещение утомляет не только зрение, но и вызывает утомление организма в целом. Неправильное освещение может быть причиной травматизма: плохо освещенные опасные зоны, слепящие лампы, резкие тени ухудшают или вызывают полную потерю зрения, ориентации. Правильное освещение уменьшает количество несчастных случаев, повышает производительность труда.

Расчёт общего равномерного искусственного освещения горизонтальной рабочей поверхности выполняется методом коэффициента использования светового потока, учитывающим световой поток, отражённый от потолка и стен. Длина помещения $A = 8$ м, ширина $B = 6$ м, высота $H = 3,5$ м. Высота рабочей поверхности над полом $h_p = 0,8$ м.

Площадь помещения:

$$S = A \cdot B = 6 \cdot 8 = 48 \text{ м}^2, \quad (1.1)$$

Коэффициент отражения стен оклеенных светлыми обоями с окнами, без штор $\rho_c = 30\%$, потолка светлой поверхности $\rho_{\Pi} = 50\%$. Коэффициент запаса, учитывающий загрязнение светильника, для помещений с малым выделением пыли равен $K_3 = 1,5$. Коэффициент неравномерности для люминесцентных ламп $Z = 1,1$.

Выбираем лампу дневного света, световой поток которой равен $\Phi_{\Pi} = 2300$ Лм.

Выбираем светильники с люминесцентными лампами. Этот светильник имеет две лампы мощностью 40 Вт каждая, длина светильника равна $A_{\text{св}} = 1227$ мм, ширина $B_{\text{св}} = 265$ мм.

Интегральным критерием оптимальности расположения светильников является величина λ , которая для люминесцентных светильников с защитной решёткой лежит в диапазоне 1,1–1,3. Принимаем $\lambda = 1,2$.

Расстояние светильников от перекрытия: $h_c = 0,4$ м.

Высота светильника над рабочей поверхностью определяется по формуле:

$$h = H - h_p - h_c = 3,5 - 0,8 - 0,4 = 2,3 \text{ м}, \quad (1.2)$$

Индекс помещения определяется по формуле:

$$i = \frac{A \cdot B}{h \cdot (A + B)} = \frac{8 \cdot 6}{2,3 \cdot (8 + 6)} = 1,49 \quad (1.3)$$

Коэффициент использования светового потока, показывающий какая часть светового потока ламп попадает на рабочую поверхность, для светильников типа ОДОР с люминесцентными лампами при $\rho_c = 30\%$, $\rho_{\Pi} = 50\%$ и индексе помещения $i = 1,49$ равен $\eta = 0,43$.

Потребные группы люминесцентных ламп светильника рассчитываются по формуле:

$$N = \frac{E \cdot A \cdot B \cdot K_3 \cdot Z}{\Phi_{\Pi} \cdot \eta} = \frac{200 \cdot 8 \cdot 6 \cdot 1,5 \cdot 1,1}{2300 \cdot 0,46} \approx 15 = 16 \quad (1.4)$$

Количество рядов светильников N_A , и количество светильников в ряду N_B рассчитывается согласно формулам 1.5-1.6. Общее число светильников: $N_{\text{св}} = 8$.

$$N_A = \frac{(B - x)}{L_2} = \frac{6000 - 709}{2128} \approx 2 \quad (1.5)$$

$$N_B = \frac{(A - y)}{L_2} = \frac{8000 - 630}{1892} \approx 4 \quad (1.6)$$

План помещения и размещения светильников с люминесцентными лампами представлен на рисунке 4.3.

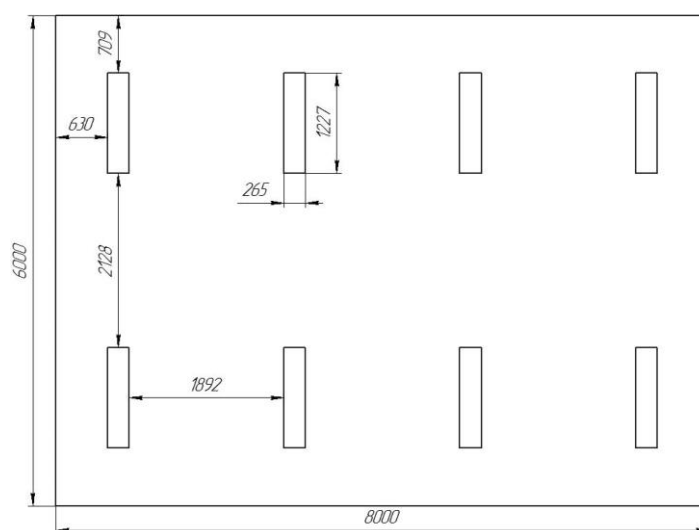


Рисунок 4.3 – План помещения и размещения светильников с люминесцентными лампами

Расчёт светового потока группы люминесцентных ламп светильника определяется по формуле:

$$\Phi_{\text{рас}} = \frac{E \cdot A \cdot B \cdot K_3 \cdot Z}{N \cdot \eta} = \frac{200 \cdot 8 \cdot 6 \cdot 1,5 \cdot 1,1}{16 \cdot 0,46} = 2152 \text{ лм} \quad (1.7)$$

Проверяем выполнения условия:

$$-10\% \leq \frac{\Phi_{\text{П}} - \Phi_{\text{рас}}}{\Phi_{\text{ЛД}}} \cdot 100\% \leq 20\%; \quad (1.8)$$

$$\frac{\Phi_{\text{П}} - \Phi_{\text{рас}}}{\Phi_{\text{ЛД}}} \cdot 100\% = \frac{2300 - 2152}{2300} \cdot 100\% = 6,4\%, \quad (1.9)$$

Исходя из полученного результата: $-10\% \leq 6,4\% \leq 20\%$, можно сделать вывод, что необходимый световой поток светильника не выходит за пределы требуемого диапазона.

5.2.3. Производственные шумы

Шум – это совокупность звуков, неблагоприятно воздействующих на организм человека и мешающих его работе и отдыху.

Допустимый уровень шума – это уровень, который не вызывает у человека беспокойства и значительных изменений показателей функционального состояния систем и анализаторов, чувствительных к шуму.

Ненормированный показатель шума на рабочих местах оказывает негативное воздействие на психологическое состояние сотрудника. У работника на поставленной ему задаче понижаются концентрация и сосредоточенность, а увеличивается утомляемость и стресс. Повышенный уровень шума приводит к нарушению слуха или являться помехой для коммуникаций между сотрудниками [29].

Измерение уровня звука и уровней звукового давления производится на расстоянии 50 см от поверхности оборудования и на уровне расположения источника(ков) звука. Уровень шума исправного компьютера находится в пределах 35-50 дБА, что значительно ниже, чем допустимый уровень шума для данного рабочего места. На рабочем месте уровень шума считается допустимым, так как не вызывает беспокойства, и не воздействует на психологическое состояние.

5.2.4. Психофизиологические факторы

При работе с ПЭВМ основным фактором, влияющим на нервную систему специалиста, является огромное количество информации, которое он должен воспринимать. Меры, позволяющие снизить воздействие этого вредного производственного фактора, регулируются СанПиН 1.2.3685-21 [26].

Для снижения влияния вредных и опасных психофизиологических факторов, влияющих на человека, основная деятельность которого предполагает работу за персональным компьютером:

- обеспечение оптимального режима труда и отдыха в течение рабочего дня;
- назначение частых, но коротких перерывов.

В частности, соблюдение данных требований позволит минимизировать нервнопсихические, нервно-эмоциональные перегрузки, а также утомление глаз, повышенную нагрузку на зрение

5.2.5. Опасность поражения электрическим током

Поражение электрическим током является опасным производственным фактором и, поскольку оператор ПЭВМ имеет дело с электрооборудованием, то вопросам электробезопасности на его рабочем месте должно уделяться много внимания. Нормы электробезопасности на рабочем месте регламентируются СанПиН 1.2.3685-21 [26], вопросы требований к защите от поражения электрическим током освещены в ГОСТ 12.1.038-82 [32].

Электробезопасность – система организационных и технических мероприятий и средств, обеспечивающих защиту людей от вредного и опасного воздействия электрического тока, электрической дуги, электромагнитного поля и статического электричества. Помещение, где расположено рабочее место оператора ПЭВМ, относится к помещениям без повышенной опасности ввиду отсутствия следующих факторов: сырость, токопроводящая пыль,

токопроводящие полы, высокая температура, возможность одновременного прикосновения человека к имеющим соединение с землей металлоконструкциям зданий, технологическим аппаратам, механизмам и металлическим корпусам электрооборудования.

Для оператора ПЭВМ при работе с электрическим оборудованием обязательны следующие меры предосторожности:

- Перед началом работы нужно убедиться, что выключатели и розетка закреплены и не имеют оголённых токоведущих частей;
- При обнаружении неисправности оборудования и приборов необходимо, не делая никаких самостоятельных исправлений, сообщить человеку, ответственному за оборудование.

5.2.6. Опасность возникновения пожара

Возникновение пожара является опасным производственным фактором, т.к. пожар на предприятии наносит большой материальный ущерб, а также часто сопровождается травмами и несчастными случаями. Регулирование пожаробезопасности производится ГОСТ 12.1.004-91 Система стандартов безопасности труда (ССБТ). Пожарная безопасность. Общие требования [33].

В помещениях с ПЭВМ повышен риск возникновения пожара из-за присутствия множества факторов:

- наличие большого количества электронных схем,
- устройств электропитания,
- устройств кондиционирования воздуха;
- возможные неисправности электрооборудования, освещения, или неправильная их эксплуатация может послужить причиной пожара.

Возможные виды источников воспламенения:

- Искры от электрооборудования;
- Искры от удара и трения;

- Открытое пламя.

5.3. Мероприятия и по снижению воздействия вредных производственных факторов

Для поддержания нормальных значений параметров микроклимата на рабочих местах рекомендуется оснащать их системами отопления, вентиляции и кондиционирования воздуха. Также, в некоторых случаях, целесообразно обеспечить питьевое водоснабжение. В помещениях для работы с ПЭВМ должна производиться ежедневная влажная уборка, а также систематическое проветривание после каждого часа работы. Для защиты операторов ПЭВМ от негативного воздействия электромагнитных полей в первую очередь необходимо, чтобы используемая техника удовлетворяла нормам и правилам сертификации.

Для создания и поддержания благоприятных условий освещения для операторов ПЭВМ, их рабочие места должны соответствовать санитарно-эпидемиологическим правилам СанПиН 1.2.3685-21 [36].

Рабочее помещение должно иметь естественное и искусственное освещение, соответствующее показателям. В качестве источников искусственного освещения должны быть использованы люминесцентные лампы, лампы накаливания – для местного освещения.

В случаях, когда характер работы требует постоянного взаимодействия с компьютером (работа программиста разработчика) с напряжением внимания и сосредоточенности, при исключении возможности периодического переключения на другие виды трудовой деятельности, не связанные с ПЭВМ, рекомендуется организация перерывов на 10–15 мин. через каждые 45–60 мин. работы. К мероприятиям по предотвращению возможности поражения электрическим током относятся:

Для предотвращения пожара помещение с ПЭВМ должно быть оборудовано первичными средствами пожаротушения: углекислотными

огнетушителями типа ОУ-2 или ОУ-5; пожарной сигнализацией, а также, в некоторых случаях, автоматической установкой объемного газового пожаротушения.

5.4. Экологическая безопасность

В данном разделе рассматривается влияние на окружающую среду деятельности по разработке проекта, а также самого продукта в результате его реализации на производстве.

К отходам, производимым в помещении, можно отнести, в первую очередь, это бумажные отходы – макулатура, пластиковые отходы, неисправные детали персональных компьютеров и других видов ЭВМ. Бумажные отходы рекомендуется накапливать и передавать их в пункты приема макулатуры для дальнейшей переработки. Снизить потребление и, соответственно, отходы бумаги можно с помощью перехода с бумажных на цифровые носители, введением цифрового документооборота и систем САПР. Также существуют программы по утилизации и переработке макулатуры, согласно ГОСТ Р 53692-2009 [35] в новый бумажный продукт, что снижает потребность в вырубке лесов. При правильной переработке более 75% макулатуры попадает во вторичный оборот. Пластиковые бутылки складывать в специально предназначенные контейнеры.

Современные ПЭВМ производят практически без использования вредных веществ, опасных для человека и окружающей среды. Исключением являются аккумуляторные батареи компьютеров и мобильных устройств. В аккумуляторах содержатся тяжелые металлы, кислоты и щелочи, которые могут наносить ущерб окружающей среде, попадая в гидросферу и литосферу, если они были неправильно утилизированы. Для утилизации аккумуляторов необходимо обращаться в специальные организации, специализировано занимающиеся приемом, утилизацией и переработкой аккумуляторных батарей.

Ртуть содержащие лампы необходимо герметично упаковать и отправить в организацию, занимающуюся их утилизацией [38]. Юридические лица обязаны сдавать лампы на переработку и вести паспорт для данного вида отходов. Дополнительным методом снижения отходов является увеличение доли электронного документооборота.

5.5.Безопасность в чрезвычайных ситуациях

Наиболее характерной для объекта, где размещаются рабочие помещения, оборудованные ПЭВМ, чрезвычайной ситуацией является пожар.

Согласно ГОСТ 12.1.004-91 [33] лаборатория относится к категории Д - негорючие вещества и материалы в холодном состоянии. Возникновение пожара при работе с электронной аппаратурой может быть по причинам как электрического, так и неэлектрического характера.

Все сотрудники организации обязаны быть ознакомлены с инструкцией по пожарной безопасности, пройти инструктаж по технике безопасности и строго соблюдать его. Запрещается использовать электроприборы в условиях, не соответствующих требованиям инструкций изготовителей, или имеющие различного рода неисправности, которые в соответствии с инструкцией по эксплуатации могут привести к пожару, а также использовать электропровода и кабели с поврежденной или потерявшей защитные свойства изоляцией.

Перед уходом из служебного помещения требуется провести его осмотр, закрыть окна, убедиться в том, что в помещении отсутствуют источники возможного возгорания, все электроприборы отключены и выключено освещение.

С периодичностью не менее одного раза в три года необходимо проводить замеры сопротивления изоляции токоведущих частей силового и осветительного оборудования. Увеличение устойчивости достигается за счет проведения соответствующих организационно-технических мероприятий, подготовки персонала к работе в ЧС.

Сотрудник при обнаружении пожара или признаков горения (задымление, запах гари, увеличение температуры и т.п.) должен:

Требуется прекратить работу, вызвать пожарную охрану по телефону «01»;

- По возможности принять меры по эвакуации людей и материальных ценностей;
- Отключить от сети электрооборудование;
- Приступить к тушению пожара имеющимися средствами пожаротушения;
- Сообщить непосредственному или вышестоящему начальнику и оповестить окружающих сотрудников;
- При общем сигнале опасности покинуть здание согласно «Плану эвакуации людей при пожаре и других ЧС».

Для тушения пожара применять ручные углекислотные огнетушители (типа ОУ-2, ОУ-5), находящиеся в помещениях офиса, и пожарный кран внутреннего противопожарного водопровода.

Рабочее место должно соответствовать требованиям пожарной безопасности, а именно, наличие охранно-пожарной сигнализации, плана эвакуации, порошковых или углекислотных огнетушителей с поверенным клеймом, табличек с указанием направления к запасному выходу (рисунок 7.4).



Рисунок 7.4 – План эвакуации.

5.6. Выводы по разделу «Социальная ответственность»

Каждый сотрудник, который работает должен проводить профессиональную деятельность с учетом социальных, правовых, экологических и культурных аспектов, вопросов здоровья и безопасности, нести социальную ответственность за принимаемые решения, осознавать необходимость устойчивого развития.

В данном разделе были рассмотрены основные вопросы соблюдения прав работника на труд, выполнения правил к безопасности труда, промышленной безопасности, экологии и ресурсосбережения.

Установлено, что рабочее место исследователя удовлетворяет требованиям безопасности и гигиены труда во время реализации проекта, а вредное воздействие объекта исследования на окружающую среду не превышает норму.

Заключение

В результате данной работы были проведены исследования различных алгоритмов обнаружения лиц и распознавание объектов. Был проведен анализ актуальности разрабатываемого проекта, исследованы и изучены различные эффективные методы и подходы, позволяющие определить правильность получаемых результатов, а также была спроектирована система распознавания лиц и обнаружения защитной маски на лице в видеопотоке. Данная система использует в своей основе методы компьютерного зрения и машинного обучения. Была рассмотрена проблема обнаружения масок, актуальная для приложений мониторинга пандемии COVID-19. Изучены недостатки набора данных для обнаружения масок, а также проведены эксперименты, при проведении которых были рассмотрены: эффективность существующих методов детектирования лиц в медицинских масках, реализации методов распознавания, направленных на обнаружение правильно надетых масок, качество работы существующих моделей обнаружения масок позволяющих производить мониторинг в задаче борьбы с COVID-19.

Разработанная система может использоваться при решении различных задач видео аналитики и имеет практическое применение в системах контроля доступа и идентификации личности, а также для выявления лиц, нарушающих санитарно-эпидемиологические требования.

Таким образом, поставленные задачи были полностью решены. Путем расширения обучающей выборки более качественными изображениями и увеличением времени обучения сверточной нейронной сети, можно добиться высокой точности работы алгоритма. В дальнейшем планируется оптимизации модели, в том числе путем изменения ее архитектуры.

Список использованных источников

1. Создать набор данных для машинного обучения [Электронный ресурс]. – Режим доступа: <https://www.machinelearningmastery.ru/how-to-build-a-data-set-for-your-machine-learning-project-5b3b871881ac/> (дата обращения: 13.11.2020).
2. Computer Vision: Algorithms and Applications (September 3, 2010 draft) 2010 г – 979 с.
3. В.Т. Фисенко, Т.Ю. Фисенко, Компьютерная обработка и распознавание изображений: учеб. пособие. - СПб: СПбГУ ИТМО, 2008. – 192 с.
4. Глубокое обучение на Python. — СПб.: Питер, 2018. — 400 с.: ил. — (Серия «Библиотека программиста»).
5. Актуальность 2D алгоритмов в определенных задачах автоматического распознавания [Электронный ресурс]. // Реноме. – Режим доступа: <https://moluch.ru/conf/tech/archive/2/138/> (дата обращения: 19.11.2020).
6. Руководство по Tensorflow: тензоры и операции [Электронный ресурс]. // – Режим доступа: <https://www.tensorflow.org/tutorials?hl=ru> (дата обращения: 14.10.2020).
7. Нейронные сети: Полный курс, 2-е издание. : Пер. с англ. – М. : Издательский дом «Вильямс», 2006. 1104 с. : ил. – Парал. тит. англ.
8. Jones, M., Viola, P. (2001) Robust Real-Time Face Detection. URL: <http://www.vision.caltech.edu/html-files/EE148-2005-Spring/pprs/viola04ijcv.pdf> (дата обращения: 8.04.2021)
9. Open Source Computer Vision [Электронный ресурс]. // – Режим доступа: https://docs.opencv.org/master/d6/d00/tutorial_py_root.html (дата обращения: 14.03.2020).
10. Мищенко Е.С. Сравнительный анализ алгоритмов распознавания лиц // Вестник Волгоградского государственного университета. Серия 9, Исследования молодых ученых. – 2013. – № 11. – С. 74–76.

11. Wang, Z.; Wang, G.; Huang, B.; Xiong, Z.; Hong, Q.; Wu, H.; Yi, P.; Jiang, K.; Wang, N.; Pei, Y.; et al. Masked face recognition dataset and application. arXiv 2020, arXiv:2003.09093. (дата обращения: 4.04.2021)
12. Rosebrock, A. COVID-19: Face Mask Detector with OpenCV, Keras/TensorFlow, and Deep Learning. [Электронный ресурс]. // – Режим доступа: <https://www.pyimagesearch.com/2020/05/04/covid-19-face-mask-detector-with-opencv-keras-tensorflow-and-deep-learning> (дата обращения: 17.11.2020).
13. Mohammad Mahdi Deramgozin, Smr. Hashemi, Azam Bastan Fard and Mohsen HajiGhorbani, "Face Recognition Improvement in Angled Status Using Invasive Weed Optimization Algorithm And fuzzy System", 2016 1st International Conference on New Research Achievements in Electrical and Computer Engineering, 2016. (дата обращения: 12.04.2021).
14. Применение масок в условиях COVID-19 [Электронный ресурс]. // – Режим доступа: https://apps.who.int/iris/bitstream/handle/10665/337199/WHO-2019-nCov-IPC_Masks-2020.5-rus.pdf (дата обращения: 17.04.2021).
15. Разработка мультимедийных приложений с использованием библиотек OpenCV и IPP [Электронный ресурс]/ А.В. Бовырин [и др.].— Электрон. текстовые данные.— М.: Интернет-Университет Информационных Технологий (ИНТУИТ), 2016.— 515 с.— Режим доступа: <http://www.iprbookshop.ru/39564> (дата обращения: 02.03.2021).
16. Оценка классификатора [Электронный ресурс]. // – Режим доступа: <http://bazhenov.me/blog/2012/07/21/classification-performance-evaluation.html> (дата обращения: 15.03.2019).
17. Буй Т.Т.Ч., Фан Н.Х., Спицын В.Г. Распознавание лиц на основе применения метода Виолы–Джонса, вейвлет-преобразования и метода

- главных компонент // Известия Томского политехнического университета. – 2012. – Т. 320, № 5. – С. 54–59.
18. Tensorflow:[Электронныйресурс] // – Режим доступа: – URL:<https://www.tensorflow.org/> (Дата обращения: 25.04.2021).
19. NumPy [Электронный ресурс] // – Режим доступа: <https://ru.wikipedia.org/wiki/NumPy> - (Дата обращения: 25.04.2021).
20. Kabir, H. M., Abdar, M., Jalali, S. M. J., Khosravi, A., Atiya, A. F., Nahavandi, S., & Srinivasan, D. (2020). Spinalnet: Deep neural network with gradual input. arXiv preprint arXiv: 2007.03347.
21. Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ (ред. от 09.03.2021).
22. ГОСТ 12.2.032-78 ССБТ «Рабочее место при выполнении работ сидя. Общие эргономические требования».
23. ГОСТ 22269-76. Система "Человек-машина". Рабочее место оператора. Взаимное расположение элементов рабочего места. Общие эргономические требования.
24. ГОСТ 12.0.003-2015 Опасные и вредные производственные факторы. Классификация. Перечень опасных и вредных факторов.
25. СанПиН 2.2.4.548-96 Гигиенические требования к микроклимату производственных помещений.
26. СанПиН 1.2.3685-21 Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания.
27. СП 52.13330.2016 Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95.
28. ГОСТ 12.1.003-2014 Система стандартов безопасности труда (ССБТ). Шум. Общие требования безопасности.

29. СН 2.2.4/2.1.8.562-96. Шум на рабочих местах, в помещениях жилых, общественных зданий и на территории жилой застройки.
30. ГОСТ 12.1.030-81 Система стандартов безопасности труда (ССБТ). Электробезопасность. Защитное заземление. Зануление.
31. ГОСТ 26522-85 Короткие замыкания в электроустановках.
32. ГОСТ 12.1.038-82 Система стандартов безопасности труда (ССБТ). Электробезопасность. Предельно допустимые значения напряжений прикосновения и токов.
33. ГОСТ 12.1.004-91 Система стандартов безопасности труда (ССБТ). Пожарная безопасность. Общие требования.
34. ГОСТ 17.4.3.04-85 Охрана природы (ССОП). Почвы. Общие требования к контролю и охране от загрязнения.
35. ГОСТ Р 53692-2009 Ресурсосбережение. Обращение с отходами. Этапы технологического цикла отходов.
36. НПБ 105-03 Определение категорий помещений, зданий и наружных установок по взрывопожарной и пожарной опасности.
37. ТОИ Р-45-084-01 Типовая инструкция по охране труда при работе на персональном компьютере.
38. ГОСТ 12.3.031-83 Система стандартов безопасности труда (ССБТ). Работы со ртутью. Требования безопасности.

Приложение А

Artificial intelligence and computer vision for object detection

Студент:

Группа	ФИО	Подпись	Дата
8BM93	Султонов Ансорхон Азизхонович		

Руководитель ВКР:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Друки А.А.	К.Т.Н.		

Консультант-лингвист отделения иностранных языков ШБИП:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Розанова Я.В.			

1.1.Machine learning

Machine learning is a collection of algorithms and tools that help machines understand patterns within data and use this underlying structure to perform reasoning about a given task. There are many ways that machines aim to understand these underlying patterns. But how does machine learning relate to deep learning? In this article, we provide an overview of how deep learning fits into this realm and also discuss some of its applications and challenges.

Machine learning (ML) is the study of computer algorithms that improve automatically through experience and by the use of data [1]. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so [2]. Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

The term machine learning was coined in 1959 by Arthur Samuel, an American IBM and pioneer in the field of computer gaming and artificial intelligence [8][9]. A representative book of the machine learning research during the 1960s was the Nilsson's book on Learning Machines, dealing mostly with machine learning for pattern classification [10]. Interest related to pattern recognition continued into the 1970s, as described by Duda and Hart in 1973[11]. In 1981 a report was given on using teaching strategies so that a neural network learns to recognize 40 characters (26 letters, 10 digits, and 4 special symbols) from a computer terminal.

1.2.Deep learning

Deep learning is a branch of machine learning which is completely based on artificial neural networks, as neural network is going to mimic the human brain so deep learning is also a kind of mimic of human brain. In deep learning, we don't

need to explicitly program everything. The concept of deep learning is not new. It has been around for a couple of years now. it gains increasing popularity due to the fact that earlier we did not have that much processing power and a lot of data. As in the last 20 years, the processing power increases exponentially, deep learning and machine learning came in the picture. A formal definition of deep learning is- neurons.

Deep learning has become widespread recently, but attempts to train algorithms based on multilayer neural networks have been carried out since the 60s of the 20th century [2]. In 1980, Kunihiro Fukushima introduced the Neocognitron model, designed to recognize patterns on the images [3].

Artificial intelligence, particularly deep learning, represents a considerable and, quite possibly, disruptive leap forward in improving the technologies supporting healthcare. Such computer algorithms are tools that can be trained to improve patient care quality by increasing diagnosis accuracy and decreasing workload stress on human care providers. They offer opportunities for automation and prediction not previously seen in healthcare. Very large data sets can be used to find correlations and patterns in clinical data in a way that is probably impossible using the unaided human brain or traditional approaches. Finally, deep learning systems can help care organizations to engage more effectively with staff and patients.

In human brain approximately 100 billion neurons all together this is a picture of an individual neuron and each neuron is connected through thousand of their neighbours. The question here is how do we recreate these neurons in a computer. So, we create an artificial structure called an artificial neural net where we have nodes or neurons. We have some neurons for input value and some for output value and in between, there may be lots of neurons interconnected in the hidden layer.

Architectures:

Deep Neural Network – It is a neural network with a certain level of complexity (having multiple hidden layers in between input and output layers). They are capable of modeling and processing non-linear relationships.

Deep Belief Network(DBN) – It is a class of Deep Neural Network. It is multi-layer belief networks.

Steps for performing DBN:

- Learn a layer of features from visible units using Contrastive Divergence algorithm.
- Treat activations of previously trained features as visible units and then learn features of features.
- Finally, the whole DBN is trained when the learning for the final hidden layer is achieved.

Recurrent (perform same task for every element of a sequence) Neural Network – Allows for parallel and sequential computation. Similar to the human brain (large feedback network of connected neurons). They are able to remember important things about the input they received and hence enables them to be more precise.

1.3. Neural networks overfitting and methods for preventing overfitting

One of the challenges for deep learning methods is overfitting. In machine learning, overfitting is the phenomenon when a trained model does not work well with examples that were not in the training dataset. Overfitting occurs when a model is trained on hidden random patterns in training data, but not in the general population. Overfitting characterizes the generalizing ability of the model.

The following regularization techniques are used for solving the problem of

overfitting:

1. L1- and L2-regularization of weights
2. Early stopping
3. The dropout method

The meaning of the first technique is limiting the absolute values of the weights. For this, one of the regularizers is applied to the loss function:

1. L1-regularizer – the sum of the squares of the weights $L1 = \lambda \sum_{\omega} \omega^2$
2. L2-regularizer – the sum of the modules of the weights $L2 = \lambda \sum_{\omega} |\omega|$

The second technique is quite simple. We should put part of the training data into the validation data. When the error begins to grow, it means that the neural network has begun to overfit and it is necessary to stop the training process. There is evidence that this technique is similar to L2-regularization [7]. Moreover, error analysis on the validation dataset allows us to evaluate the generalized ability of models.

The third method is more effective than the first. The idea of the “dropout” method is that for each neuron (except for the output layer), the probability p is established with which it will be ejected from the neural network. The training algorithm changes in this way: on each new training example x , we first determine for each neuron activity with probability p , and depending on the result we either use the neuron as usual, or set its output to always be strictly equal to zero. Then everything happens without changes. Zero at the output leads to the fact that the neuron actually

1.4.Convolutional neural networks

A convolutional neural network (also CNN or ConvNet) is one of the most

popular algorithms in deep learning, a type of machine learning in which a model learns to perform classification tasks directly on an image, video, text, or sound.

Convolutional neural networks are especially useful for finding patterns in images for recognizing objects, faces, and scenes. They learn directly from images, using templates for image classification and eliminating the manual feature extraction.

Applications requiring object recognition and computer vision, such as self-driving cars and facial recognition, rely heavily on CNN. Depending on your application, you can build your CNN from scratch, or use a ready-made model based on your data.

The main idea of convolutional neural networks is: the processing of image part should not depend on the specific location of this part. Therefore, we cover the input image with small windows (for example, 5×5 pixels) and extract features in each window with a small neural network. It will have only $5 \times 5 = 25$ inputs, and from each image it can get a lot of different inputs [8]. Then the results can again be presented in the form of an image, replacing the 5×5 windows with their central pixels. Then we can apply a second convolutional layer with another small neural network to this image. In each convolutional layer there are very few free parameters, especially in comparison with fully connected neural networks [8].

Typically, the original image consists of three matrices (by RGB color channels) – red, green and blue. These matrices are called channels. The same matrices will be obtained after each convolutional layer. They have a spatial structure corresponding to the original picture. There can now be more channels. The values of each feature that we extracted from the windows in the original image will now be a whole matrix. Each matrix is called a feature map [8].

Next, we will consider how convolutional layers work. A convolution is a

special linear transformation of input data. If x^l is the feature map in the layer with number l , then the result of 2d-convolution with a kernel of size $2d + 1$ and a weight matrix W of size $(2d + 1) \times (2d + 1)$ on the next layer will be like this [10]:

$$y_{i,j}^l = \sum_{-d \leq a, b \leq d} W_{a,b} x_{i+a, j+b}^l$$

where $y_{i,j}^l$ – result of convolution on layer l , $x_{i,j}^l$ – its input, i.e. the output of the entire previous layer. We apply a linear transformation to the square window of the previous level to obtain a component of the next level, that is, we scalarly multiply the pixels from the window by the convolution vector.

The convolution operation has the following properties:

1. The convolution saves the structure of the input image, as it is applied to each section of the input data separately;
2. The convolution operation has the sparse property, since the value of each neuron of the next layer depends only on a small fraction of the input neurons;
3. The convolution uses the same weights many times, as they are reapplied to different parts of the entrance.

We can apply the nonlinear function h after the convolution operation [4]. It will simply be applied to each element of the resulting tensor individually.

In the classical convolutional layer, in addition to linear convolution and non-linearity following it, there is one more operation: pooling. In convolutional networks, the fact of the presence or absence of one or another feature is usually more important than its exact coordinates. For example, in the case of face recognition, it is more important to understand whether the image has a face and whose face it is than to find out where it is located. Therefore, we can generalize the

extracted features, having lost some of the information about their location, but at the same time reducing the dimension [6].

Usually max-pooling operation is applied to each local group of neurons. The formula for calculating pooling is as follows [10]:

$$x_{i,j}^{l+1} = \max_{-d \leq a \leq d, -d \leq b \leq d} z_{i+a,j+b}^l$$

An example of the operation of the pooling layer is shown in Figure 1.

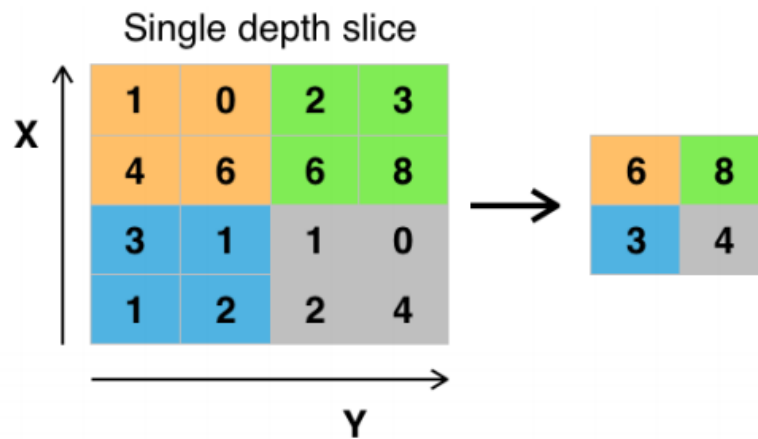


Fig. 1 Max-pooling

Another important element is the activation function, which was mentioned above. The activation function after the convolution layer is a source of non-linearity, which allows us to train neural networks. The activation function must meet the following criteria:

1. To be monotonously non-decreasing;
2. To be limited, i.e. tend to zero at $x \rightarrow -\infty$;
3. To be differentiable.

There are many activation functions: logistic sigmoid, hyperbolic tangent, and others. However, rectified linear unit (ReLU) is more popular function for convolutional neural networks [10]. It has the following formula [10]:

$$ReLU(x) = \begin{cases} 0, & \text{если } x < 0, \\ x, & \text{если } x \geq 0. \end{cases}$$

ReLU is more effective than logistic sigmoid and hyperbolic tangent, since complex calculations are not needed to calculate the derivative of this function: if $x < 0$, then the derivative is zero, if $x \geq 0$, then the derivative is unity. It also allows you to reduce the cost of computing resources. In addition to the classic ReLU implementation, there are modifications of it that allow you to get the best result in some tasks [6].

1.5.Face recognition systems

Face recognition is a programm algorithm used to verify or identify a person's identity by processing a video frame or digital image that shows a person's face. Facial recognition technologies use several different methods, but they usually compare the facial features in the image with the faces contained in the database.

A face recognition process is a collection of different tasks that serve to identify a face from a digital image or video. In general terms, this process occurs as follows steps: after the system has received an image from the camera, the first step in the system is face detection, which determines whether the image or video contains any faces. Face detection also locates key features (called “landmarks”) such as eyes, nose, etc. These landmarks also can be used to locate a face in an image or video and determine its size and orientation. Further, using algorithms, the boundaries of the face are determined (detection stage). Accordingly, the stage of recognition follows, in which the face is transformed (its brightness changes, it is aligned, scaled, etc.) and brought to a certain predetermined form. At the final stage, features are detected and directly compared with features from databases. This stage of comparison is called identification or verification.

Verification: This is a one-to-one pattern matching process (1:1). To determine an identity, the system compares a biometric sample with one biometric

template stored in a database to predict if they are the same.

Identification: This is a one-to-many (1: N) matching process - where the system compares the face in the image and the biometric sample to all face templates stored in the database.

Figure 2 (look below) shows the general algorithm for face recognition from an image or video.

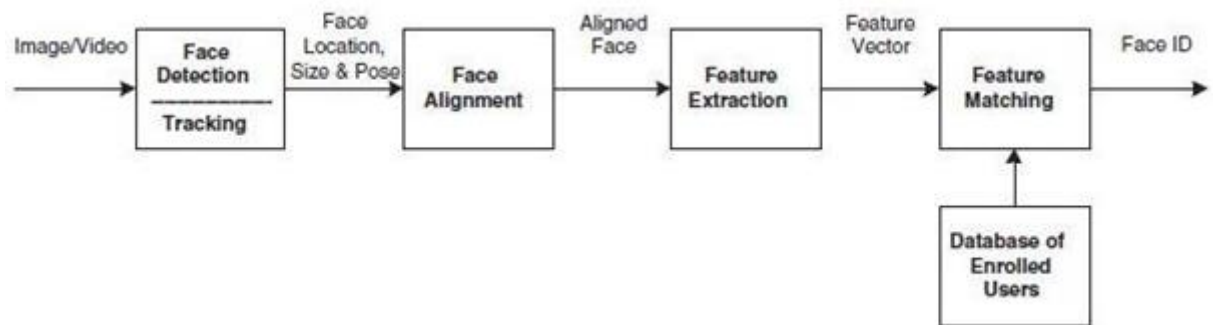


Fig. 2 Face recognition algorithm

1.6.Face detection methods

There are many face detection algorithms. Most of them are available as articles describing mathematical models, while others are presented as software implementations. An image in the form of digital data from cameras is transmitted to a computer - it is processed using a special algorithm that determines the location of the face area by its main features (eyes, mouth, eyebrows, nose, etc.). There are many similar methods for detecting faces, and most of them are a combination of other methods.

This section provides a classification of face detection algorithms and describes the advantages and disadvantages of each group of algorithms. But all of them can be divided into two categories: methods based on knowledge and methods

of facial detection based on external signs. The classification of detection methods is shown in Figure 3 (look below).

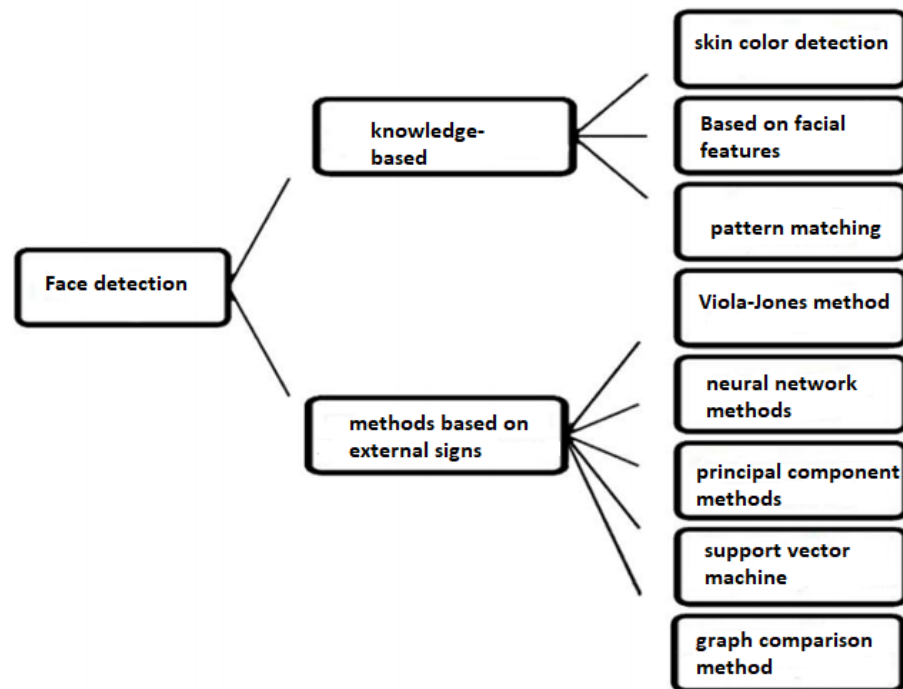


Fig. 3 Classification of face detection methods

Knowledge-based methods depend on a set of rules and it is based on human knowledge to detect faces. These methods use information about the face (the face must have a nose, eyes and mouth at a certain distance from each other), shape, skin color, texture, or its features. All rules are formalized knowledge that a person is guided by. For example, the basic rules are: the areas of the eyes, nose and mouth differ in brightness relative to the rest of the face; the eyes on the face are always located symmetrically relative to each other. The big problem with these methods is the difficulty in creating an appropriate set of rules. There can be many false positives if the rules were too general or too detailed. This approach alone is not enough and it is impossible to find many faces in multiple images.

The same group of methods includes a more general method - the method of pattern matching. The pattern matching method uses predefined or parameterized

face patterns to determine or detect faces based on the correlation between patterns and input images. For example, a human face can be divided into eyes, face contour, nose and mouth. In this method, descriptions of the properties of individual areas of the face and their specified relative position, which will help to find a match with the original image according to the templates and standard of the face.

Knowledge-based methods are fairly widespread and perform well, but they only perform well on images with good resolution, no noise, and a simple background. In frames from a video stream or cameras installed in public places, where different angles and turns of faces are possible, as well as changing lighting and a large number of interfering objects in the background, there is a high probability of errors.

The face detection method based on external features depends on the set of trainable images of the delegate's face to determine the face models. Appearance-based approach is better than other executions. In general, the outwardly based approach relies on statistical analysis and machine learning techniques to find relevant characteristics of facial images. This technique is also used for feature extraction and face recognition. Accordingly, the methods of this group do not have the above-mentioned drawbacks and they have become more often used in video surveillance systems. Face detection in such methods is performed by iterating over all fragments of the image to determine which class the image belongs to: to the class of images without a face, go to the class containing a face.

Из-за большого объема работы, методы обладают избыточностью и большой вычислительной сложностью. Для того чтобы уменьшить количество вычислений и ускорить процесс отыскания лиц, применяются дополнительные методы для сокращения количества рассматриваемых фрагментов.

The most relevant face detection methods are discussed below:

1.7.Viola-Jones method

Viola-Jones object detection method. The method was named after two computer vision researchers who proposed the method in 2001, Paula Viola and Michael Jones, in their article "Rapid Object Detection Using an Enhanced Cascade of Simple Functions." Although the Viola-Jones method is outdated, it is quite powerful and has proven to be extremely noticeable in real-time face detection. This algorithm can detect faces in real time at an impressive speed. [eight]. This method has many implementations, including as part of the OpenCV computer vision library [9].

The main stages of the method:

- images are used in an integral representation, which allows you to quickly calculate the necessary objects;
- Haar signs are used, with the help of which the search for the desired object is carried out (in this context, a person and its features);
- boosting is used (from the English boost - improvement, amplification) to select the most suitable features for the desired object in this part of the image; all signs go to the input of the classifier, which gives the result "true" or "false";
- cascades of features are used to quickly discard windows where a face is not found.

Elements of the Haar cascade are composed of dark and light areas. It produces one value by taking the sum of the intensities of the light areas and subtracting that from the sum of the intensities of the dark areas. These Haar features give a point value of the difference in brightness along the X and Y axes, respectively. Therefore, a common Haar feature for face recognition is a set of two

adjacent rectangles that lie above the eyes and on the cheeks. The characteristic value is calculated by the formula:

$$F = X - Y \quad (1)$$

Where X is the sum of the values of the brightness of the points covered by the light part of the feature, and Y is the sum of the values of the brightness of the points covered by the dark part of the feature.

The integral representation of images can be represented in the form of a matrix, the dimensions of which coincide with the dimensions of the original image I, and each element of it stores the sum of intensities of all pixels located to the left and above this element. Matrix elements are calculated using the following formula:

$$L(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j) \quad (2)$$

Where I (i, j) is the pixel brightness of the original image.

Each element of the matrix L (x, y) is the sum of pixels in a rectangle from (0,0) to (x, y), i.e. the value of each pixel (x, y) is equal to the sum of the values of all pixels to the left and above this pixel (x, y).

To calculate the sum of a rectangular area in the integral representation of an image, only 4 operations of accessing an array and 3 arithmetic operations are required. This allows you to quickly calculate the Haar features for images in training and recognition [7].

Using such an integral matrix, you can very quickly calculate the sum of the pixels of an arbitrary rectangle, an arbitrary area.

Let the rectangle ABCD contain the object D of interest to us:

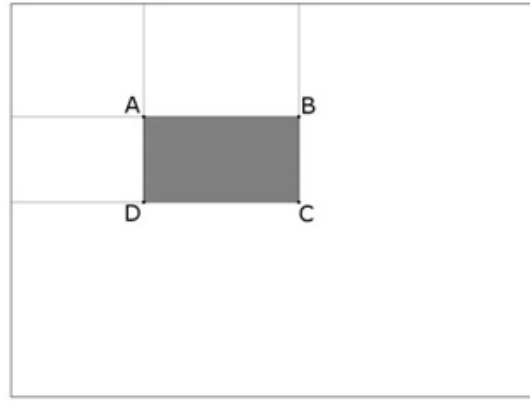


Fig. 4 Is a rectangular area using an integral image.

The sum inside the rectangle ABCD can be expressed in terms of the sums and differences of adjacent rectangles using the formula:

$$S(ABCD) = L(A) + L(C) - L(B) - L(D) \quad (3)$$

AdaBoost (Adaptive Boosting) algorithm is a machine learning algorithm for choosing the best subset of features among all available features. The result of the operation of the algorithm is a classifier (aka a forecasting function, a hypothesis function), called a "strong classifier" [7]. A strong classifier consists of linear combinations of "Weak classifiers" (best characteristics). From a high level, to find these weak classifiers, the algorithm is run for T iterations, where T is the number of weak classifiers to be found. At each iteration, the algorithm finds the error rate for all functions and then selects the function with the lowest error rate for that iteration. AdaBoost chooses a set of weak classifiers to combine and assigns its own weight to each of them. This weighted combination is a strong classifier.

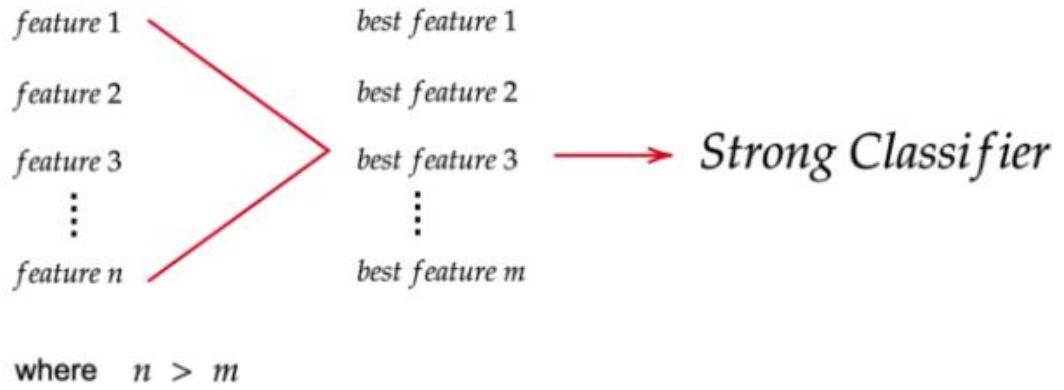


Fig. 5 Purpose of Using the AdaBoost Algorithm

For each feature, the weak classifier determines the optimal classification threshold function that provides the minimum number of misclassified examples. The weak classifier h thus consists of the values of the attribute of the object f , threshold and parity p , indicating the direction of the inequality sign.

As a result of the boosting algorithm at each iteration, a simple classifier of the form is formed:

$$h_j(z) = \begin{cases} 1, & \text{if } p_j f_j(z) < p_j Q_j \\ 0, & \text{else} \end{cases}$$

Where, p_j shows the direction of the inequality sign, Q_j is the threshold value for j , $f_j(z)$ is the calculated value of the feature, z is the image window.

A cascade classifier is a multi-stage classifier that can perform detection quickly and accurately. Each stage consists of a strong classifier generated by the AdaBoost algorithm. From stage to stage, the number of weak classifiers in a strong classifier increases [7]. Input data is evaluated sequentially (step by step). If the classifier fails for a particular stage, the input is immediately discarded. If the output is positive, the input is redirected to the next stage, as shown in Figure 6 (look below).

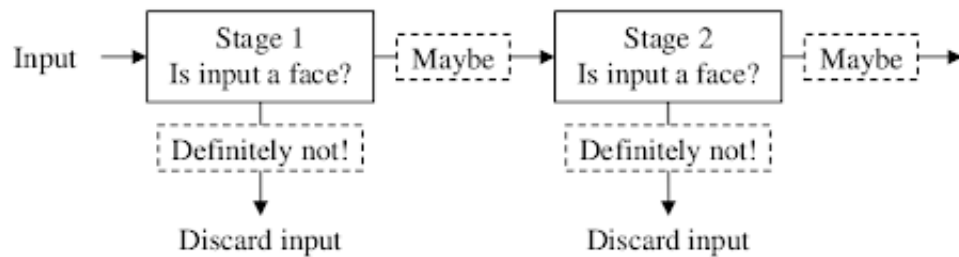


Fig. 6 Cascade classifier.

When a filter passes an area of the image, it moves to the next filter in the sequence. When the entire area of the image has gone through all the filters, then it is classified as a "face", while the steel ones are classified as "not a face".

Advantages of the Viola-Jones Method:

- Possesses high accuracy of detecting rotated faces at different angles (up to 60-65) degrees;
- Speed of work;

Disadvantages:

- A large number of closely spaced results due to the use of different scales and sliding window;
- Long learning time. The algorithm needs to analyze a large number of test images.

1.8.Face recognition methods

Another important part of automatic face recognition systems is directly the human recognition algorithms themselves. Today there are many such algorithms, and each of them has its own specifics, its own speed of operation and its own recognition reliability. Recognition algorithms are divided into two categories,

depending on the recognition technology used - two-dimensional, in which recognition is based on the geometry of the face (2D technology) and three-dimensional, in which recognition is based on the structure of the skull (3D technology).

2D recognition systems work with "flat", two-dimensional images and recognize a face by analyzing its texture and areas of the face with high contrast, so if the lighting or the position of the face is disturbed, its recognition becomes very difficult.

3D recognition systems are more resistant to such changes, because when creating a model of a face, they take into account the peculiarities of the structure of the skull. But nevertheless, although 3D systems have such a big plus, due to the lack of the ability to serve a large number of users in identification mode and due to low speed, 3D technologies have not yet been widely used. Also, 3D recognition technologies require large computing resources, and the cost of equipment for such systems is much higher than that of 2D [5].

Приложение Б. Исходный код процесса обучения и проверки моделей

```
import tensorflow as tf
from tensorflow import keras
from sklearn.preprocessing import LabelBinarizer
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report

# Вспомогательные библиотеки
import numpy as np
import matplotlib
import matplotlib.pyplot as plt

# инициализируем начальную скорость обучения, количество эпох для
тренировки,
INIT_LR = 1e-4
EPOCHS = 10
BS = 32

# список данных (т.е. изображений) и изображений классов
print("[INFO] loading images...")
imagePaths = list(paths.list_images(args["dataset"]))
data = []
labels = []

# перебираем пути к изображениям
for imagePath in imagePaths:
    # извлекаем метку класса из имени файла
    label = imagePath.split(os.path.sep)[-2]

    # загружаем исходное изображение (224x224) и обрабатываем ее
    image = load_img(imagePath, target_size=(224, 224))
    image = img_to_array(image)
    image = preprocess_input(image)
```

```

# обновить списки данных и меток соответственно
data.append(image)
labels.append(label)

# преобразовываем данные и метки в массивы с помощью NumPy
data = np.array(data, dtype="float32")
labels = np.array(labels)

model = Sequential([
    Conv2D(100, (3, 3), activation='relu', input_shape=(150, 150, 3)),
    MaxPooling2D(2, 2),

    Conv2D(100, (3, 3), activation='relu'),
    MaxPooling2D(2, 2),

    Flatten(), # Сглаживает ввод. Не влияет на размер партии.
    Dropout(0.5),
    Dense(50, activation='relu'),
    Dense(2, activation='softmax')
])

model.compile(optimizer='adam', loss='binary_crossentropy',
metrics=['acc'])

# Генерация данные изображение для обучение
TRAINING_DIR = "./train"
train_datagen = ImageDataGenerator(rescale=1.0/255,
                                    rotation_range=40,
                                    width_shift_range=0.2,
                                    height_shift_range=0.2,
                                    shear_range=0.2,
                                    zoom_range=0.2,
                                    horizontal_flip=True,
                                    fill_mode='nearest')

```

```

train_generator = train_datagen.flow_from_directory
(
    TRAINING_DIR,
    batch_size=10,
    target_size=(150, 150))

VALIDATION_DIR = "./test"
validation_datagen = ImageDataGenerator(rescale=1.0/255)

validation_generator =
validation_datagen.flow_from_directory(VALIDATION_DIR,

batch_size=10,

target_size=(150, 150))

# Создаем контрольную точку обратного вызова, чтобы сохранять лучшую
модель после каждой эпохи во время обучения:

checkpoint = ModelCheckpoint('model2-{epoch:03d}.model',
monitor='val_loss',verbose=0,save_best_only=True,mode='auto')

#Обучение модели
history = model.fit_generator(train_generator,

                                epochs=10,

                                validation_data=validation_generator,

                                callbacks=[checkpoint])

```

Визуализация результата обучения по эпохам

```

Epoch 1/10
132/132 [=====] - 171s 1s/step - loss: 0.6702 - a
cc: 0.7141 - val_loss: 0.1702 - val_acc: 0.9588
Epoch 2/10
132/132 [=====] - 167s 1s/step - loss: 0.3395 - a
cc: 0.8730 - val_loss: 0.0491 - val_acc: 0.9072

```



```

Epoch 3/10
132/132 [=====] - 166s 1s/step - loss: 0.3382 - a
cc: 0.8852 - val_loss: 0.2210 - val_acc: 0.9021
Epoch 4/10
132/132 [=====] - 166s 1s/step - loss: 0.2634 - a
cc: 0.8958 - val_loss: 0.0169 - val_acc: 0.9742
Epoch 5/10
132/132 [=====] - 171s 1s/step - loss: 0.2149 - a
cc: 0.9118 - val_loss: 0.1482 - val_acc: 0.9588
Epoch 6/10
132/132 [=====] - 166s 1s/step - loss: 0.2057 - a
cc: 0.9278 - val_loss: 0.0036 - val_acc: 0.9691
Epoch 7/10
132/132 [=====] - 166s 1s/step - loss: 0.2048 - a
cc: 0.9346 - val_loss: 0.0698 - val_acc: 0.9588
Epoch 8/10
132/132 [=====] - 166s 1s/step - loss: 0.1803 - a
cc: 0.9346 - val_loss: 0.0073 - val_acc: 0.9845
Epoch 9/10
132/132 [=====] - 178s 1s/step - loss: 0.1570 - a
cc: 0.9414 - val_loss: 0.0013 - val_acc: 0.9845
Epoch 10/10
132/132 [=====] - 167s 1s/step - loss: 0.1334 - a
cc: 0.9536 - val_loss: 0.0335 - val_acc: 0.9536

```

В процессе обучения модели отображаются метрики потери (loss) и точности (ассурацы). Эта модель достигает на тренировочных данных точности равной приблизительно 0.98 (98%).

Теперь протестируем результаты модели детектора маски на лице с помощью видеопотока получаемый из камер.

```

import cv2
import numpy as np
from keras.models import load_model
model=load_model("./model-09")

results={0:'without mask',1:'mask'}
GR_dict={0:(0,0,255),1:(0,255,0)}

rect_size = 4
cap = cv2.VideoCapture(0)

haarcascade =
cv2.CascadeClassifier('C:/Users/Ansor/anaconda3/envs/LABML/Lib/site-
packages/cv2/data/haarcascade_frontalface_default.xml')

while True:
    (rval, im) = cap.read()

```

```

im=cv2.flip(im,1,1)

rerect_size = cv2.resize(im, (im.shape[1] // rect_size,
im.shape[0] // rect_size))
faces = haarcascade.detectMultiScale(rerect_size)
for f in faces:
    (x, y, w, h) = [v * rect_size for v in f]

    face_img = im[y:y+h, x:x+w]
    rerect_sized=cv2.resize(face_img, (150,150))
    normalized=rerect_sized/255.0
    reshaped=np.reshape(normalized, (1,150,150,3))
    reshaped = np.vstack([reshaped])
    result=model.predict(reshaped)

    label=np.argmax(result,axis=1)[0]

    cv2.rectangle(im, (x,y), (x+w,y+h), GR_dict[label], 2)
    cv2.rectangle(im, (x,y-40), (x+w,y), GR_dict[label], -1)
    cv2.putText(im, results[label], (x, y-
10), cv2.FONT_HERSHEY_SIMPLEX, 0.8, (255,255,255), 2)

    cv2.imshow('LIVE', im)
    key = cv2.waitKey(10)

    if key == 27:
        break
cap.release()
cv2.destroyAllWindows()

```